Дискретная математика

КОНСПЕКТЫ ЛЕКЦИЙ
МФТИ 2022

Александр Рубцов
alex@rubtsov.su

Содержание

Введение

В рамках этого курса мы изучим базовые математические понятия, такие как множества, функции, графы, бинарные отношения. Они важны и нужны для дальнейшего изучения Computer Science, но также являются базой и для чистой математики. С другой стороны, для плодотворного развития в Computer Science нужно понимать, что такое доказательство. Доказательства корректности многих алгоритмов являются по сути доказательствами теорем, поэтому в рамках этого курса мы будем уделять доказательствам особое внимание.

Мы начинаем изучение курса с алгебры логики. С одной стороны, эта тема связана с чистым Computer Science — логика в программировании используется постоянно, хотя бы для задания условий выхода из цикла и условных операторов. С другой стороны, логические законы являются законами для доказательств теорем и для математических рассуждений в целом.

Лекция 1

Алгебра логики: введение.

План:

1. Высказывания и логические связки.
2. Булевы функции и способы их задания: таблицы истинности, вектор значений, формулы.
3. Законы коммутативности, ассоциативности и дистрибутивности, приоритет операций.
4. Законы поглощения.
5. Равенство булевых функций (и булевых формул). Существенные и фиктивные переменные.

Ключевые слова: *высказывание, булева функция, конъюнкция, дизъюнкция, импликация, отрицание (инверсия), эквивалентность, XOR (исключающее или), коммутативность, ассоциативность, дистрибутивность, законы поглощения, существенная переменная, фиктивная переменная.*

Литература: [1], [2]

Мы начинаем наш курс с основ логики, поскольку логика — это цемент для построения математических утверждений и доказательств. Хотя мы и начинаем фактически с изучения правил переписывания логических формул, которые можно выполнять не вдаваясь в суть самих высказываний, дальше эти правила будут использоваться для построения утверждений и доказательств.

Одной из основных целей нашего курса является прививание первокурсникам математической культуры. Математическая культура начинается с работы с формальными определениями, и это влечёт трудности. Формальные определения требуют язык теории множеств и часто при первом изучении формализм мешает

содержательному пониманию вещей. Мы решаем эту проблему следующим образом. До введения теории множеств все определения будут неформальными, а после её изучения мы формализуем уже введённые определения на языке теории множеств. Однако после изучения теории множеств мы не откажемся от неформальных определений перед формальными. Определений в курсе достаточно много, поэтому они часто не выделяются в отдельные блоки, а просто по ходу текста выделяется ***определяемое понятие***. Можно было бы начать изложение с теории множеств, а не логики, но тогда у нас была бы другая проблема — пришлось бы вести изложение без объяснения, что считается доказательством. Надеюсь, мы выбрали меньшее из зол.

Итак, перейдём к изучению основ логики. Алгебра логики оперирует с высказываниями. ***Высказывание*** — это утверждение, которое либо истино, либо ложно. Истинность высказывания A будем обозначать «1», а ложность — «0». В роли A может выступать высказывание «за окном идёт дождь» или «число 7 делится на 6». Заметьте, что утверждение «число x делится на 6» не является высказыванием, в случае если число x не зафиксировано; тогда утверждение зависит от параметра и в зависимости от значения x может меняться истинность этого утверждения. Оговоримся, что утверждение, зависящее от параметра, не считается высказываением даже в случае, когда оно истинно при любом значении параметра: например, «число x делится на 1» — не высказывание.

С помощью логических связок из высказываний можно получать более сложные высказывания, такие как «A и B», «A или B», «не A». Вы знакомы с этими связками со школьной скамьи и знаете, что их называют «конъюнкция», «дизъюнкция» и «отрицание» («инверсия») и обозначают

$$A \wedge B, \qquad A \vee B, \qquad \neg A \text{ или } \bar{A}.$$

Что же такое логические связки? Это функции, которые зависят от набора переменных, принимающих значения 0 или 1 (от набора высказываний). Такие переменные называют ***булевыми переменными***, а функции — ***булевыми функциями***. Есть несколько стандартных способов задания булевой функции. Мы начнём с таблицы истинности.

A	B	$A \wedge B$	$A \vee B$	$A \to B$	$A \leftrightarrow B$	$A \oplus B$
0	0	0	0	1	1	0
0	1	0	1	1	0	1
1	0	0	1	0	0	1
1	1	1	1	1	1	0

Рис. 1.1. задание логических связок таблицами истинности.

Если функция зависит от k переменных, то первые k столбцов соответствуют переменным, а $k+1$-й столбец — значению функции на соответсвующем наборе значений переменных (все наборы значений перечисляются в строках таблицы). В таблице на рис. 1.1 мы скомпоновали таблицы истинности для нескольких функций в одну таблицу.

Пожалуй, все логические связки имеют естественную интерпретацию, кроме импликации (почему её определяют именно так, мы обсудим позже):

Обозначение	**Смысл**	**Название**
$A \wedge B$	«A и B»	конъюнкция
$A \vee B$	«A или B»	дизъюнкция
$\neg A$	«не A»	отрицание
$A \to B$	«из A следует B»	импликация
$A \leftrightarrow B$	«A равносильно B»	эквивалентность
$A \oplus B$	«либо A, либо B»	XOR (исключающее или)

Наборы значений переменных принято перечислять в следующем порядке. Первым идёт набор из одних нулей, а дальше i-й набор является двоичной записью числа $i-1$. Таким образом, всего в таблице истинности 2^k строк (именно столько чисел имеют двоичную запись длины k).

Благодаря стандартному порядку можно просто задать булеву функцию столбцом её значений:

$$f(x_1) = 10 = \neg x_1, \quad g(x_1, x_2) = 0001 = x_1 \wedge x_2, \quad h(x_1, x_2) = 0110 = x_1 \oplus x_2.$$

Говорят, что функция задана ***вектором значений***.

1.1 Задание булевых функций с помощью логических связок (формулами)

Следующий способ задания булевой функции — использовать уже заданные булевы функции для определения более сложных функций. С точки зрения логики, мы используем логические связки для построения более сложных высказываний, например

$$(A \wedge B) \to C.$$

Формулу можно изобразить с помощью дерева, вычисления по которому (после присваивания значений переменным) идут снизу вверх; вообще говоря вычислять значение каждого узла дерева можно параллельно.

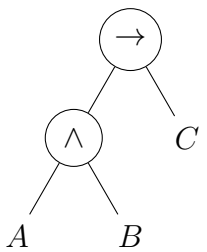


Рис. 1.2. Дерево для формулы $(A \wedge B) \to C$

Скобки используются для указания приоритета операций: в формуле выше следует сначала выполнить конъюнкцию, а затем импликацию. Для удобства записи используют следующие договорённости о приоритете операций:

1) отрицание $\neg$;

2) конъюнкция $\wedge$;

3) дизъюнкция $\vee$ и XOR $\oplus$;

4) импликация $\to$;

5) эквивалентность $\leftrightarrow$.

Согласно этим договорённостям можно восстановить скобки: самый высокий приоритет у отрицания, самый низкий — у эквивалентности. В случае, если две операции имеют одинаковый приоритет, то считается, что сперва выполняется та, которая стоит левее. Но лучше поставить лишнюю пару скобок, дабы не было непонимания, что мы и будем делать.

Пример 1. Расставив скобки для формулы $x_1 \to \neg x_2 \wedge x_1 \vee x_2$ получим

$$(x_1 \to (((\neg x_2) \wedge x_1) \vee x_2)).$$

Расставив скобки, согласно приоритету можно преобразовать формулу в дерево, руководствуясь следующими правилами. Если формула состоит только из переменной, то дерево состоит только из одного узла — самой переменной. Если в формуле есть операции, то нужно взять операцию из внешних скобках и сделать её вершиной дерева; построить деревья для левого и правого операндов и сделать их левым и правым детьми внешней операции соответственно.

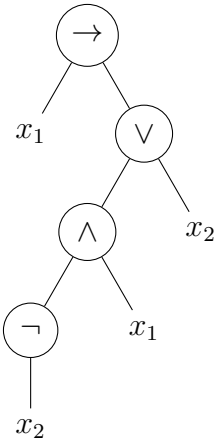


Рис. 1.3. Дерево для формулы из примера 1

1.2 Равенство булевых функций и формул. Фиктивные переменные

Один из самых коварных знаков в математике — это знак равенства. Справедливость формулы $x_1 \wedge x_2 = x_2 \wedge x_1$ не вызвает сомнений, ведь можно подставить в неё

всевозможные наборы значений переменных и проверить, что значение левой части всегда равняется значению правой. Разумно было бы сказать, что левая и правая части равенства задают булевы функции, и равенство справедливо, если булевы функции совпадают, но тут есть одна загвоздка. Рассмотрим теперь равенство

$$x_1 \wedge x_2 = x_2 \wedge x_1 \wedge (x_2 \vee \neg x_3). \tag{1}$$

Это равенство также легко проверить подстановкой, но в его левой части записана булева функция от двух аргументов, а в правой — от трёх. От этого можно было бы отмахнуться, сказав, что можно считать, что левая часть равенства задаёт функцию от трёх аргументов, но глядя только на левую часть невозможно предсказать сколько переменных понадобится. Любой, кто хоть раз пробовал писать на C++ понимает, что эта проблема довольно деликатная.

Решение этой проблемы приводит к определению равенства булевых функций, но начнём мы с другого определения. ***Тавтология*** — это булева функция, которая возвращает 1 на любом наборе переменных. Булевы функции f и g ***равны***, если функция $h = (f \leftrightarrow g)$ — тавтология. Определение равенства булевых функций переносится на равенство формул. Если не оговорено противного, то мы считаем, что формула задаёт булеву функцию, зависящую только от переменных, встречающихся в этой формуле.

Обозначим через $g(x_1, x_2, x_3)$ булеву функцию, заданную правой частью формулы (1). Ясно, что значение g не зависит от переменной x_3. Благодаря приведённому определению равенства, мы получаем, что $g(x_1, x_2, x_3) = g(x_1, x_2, 0) = g(x_1, x_2, 1)$. В случае, если для булевой функции $f(x_1, \ldots, x_n)$ справедливо равенство

$$f(x_1, x_2, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) = f(x_1, x_2, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n), \tag{2}$$

переменная x_i называется ***фиктивной*** (для функции f); в случае, если равенство (2) не выполняется для переменной x_i, то она называется ***существенной***.

Заметим, что для доказателсьства существенности переменной x_i достаточно найти в таблице истинности f два набора значений переменных, отличающиеся только значениями x_i, на которых f принимает разные значения. Для доказательства же фиктивности, необходимо проверить совпадание значений f на всех таких наборах: равенство (2) выполняется если и только если оно выполняется для всех наборов значений переменных $x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$.

1.3 Логические тождества

Для логических связок выполняется довольно много законов (тождеств). Их справедливость можно проверить с помощью таблиц истинности.

Законы коммутативности:

$$x_1 \wedge x_2 = x_2 \wedge x_1, \qquad x_1 \vee x_2 = x_2 \vee x_1, \qquad x_1 \oplus x_2 = x_2 \oplus x_1.$$

Законы ассоциативности:

$$x_1 \wedge (x_2 \wedge x_3) = (x_1 \wedge x_2) \wedge x_3.$$

Законы дистрибутивности:

- $A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$
- $A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$
- $A \vee (B \to C) = (A \vee B) \to (A \vee C)$
- $A \to (B \wedge C) = (A \to B) \wedge (A \to C)$
- $A \to (B \vee C) = (A \to B) \vee (A \to C)$
- $A \to (B \to C) = (A \to B) \to (A \to C)$

Правила поглощения:

$$A \vee (A \wedge B) = A, \quad A \wedge (A \vee B) = A.$$

Другие важные свойства:

- $A \wedge A = A, \quad A \vee A = A$ (идемпотентность)
- $A \wedge \neg A = 0, \quad A \vee \neg A = 1$ (дополнение)
- $A \wedge 0 = 0, \quad A \vee 0 = A$ (универсальные границы)
- $A \wedge 1 = A, \quad A \vee 1 = 1$ (универсальные границы)
- $\neg(\neg A) = A$ (инволютивность)
- $\neg(A \wedge B) = \neg A \vee \neg B, \neg(A \vee B) = \neg A \wedge \neg B$ (законы Моргана)
- $A \to B = \neg A \vee B, \quad A \vee B = \neg A \to B$

Просто чертить таблицы истинности для проверки данных тождеств довольно времязатратно. Приведём пример доказательства тождества через таблицу истинности, не рисуя саму таблицу.

Пример 2. Докажем тождество $A \to B = \neg A \vee B$. И импликация и дизъюнкция имеют ровно один ноль среди значений: на наборах $(1, 0)$ и $(0, 0)$ соответственно. Таким образом, заменив в дизъюнкции $A \vee B$ переменную A на её отрицание, мы получим таблицу истинности для импликации.

Многие тождества среди других важных свойств получаются путём подстановки вместо переменной констант.

Тождества, которые мы рассматривали выше, иллюстрируют свойства логических связок. Логических законов за ними сходу не видно. Не такие, например, «закон двойного отрицания» $\neg\neg A = B$ и закон контрапозиции $A \to B = \neg B \to \neg A$.

С осмысленными логическими законами мы познакомимся чуть позже, а пока учимся утилитарно работать с алгеброй логики.

Лекция 2

Множества и логика

План:

1. Множества и операции над ними
2. Связь алгебры логики и алгебры множеств
    - предикаты
    - юнивёрсум и дополнение
    - законы де Моргана
    - кванторы
    - эквивалентность тождеств алгебры множеств и алгебры логики
    - импликация
    - контрапозиция

Литература: [1], [3]

В математических курсах часто стараются, чтобы всё было достаточно строго определено. Однако, для совсем базовых понятий приходится делать исключение: точки и прямые в школьной геометрии не определяют, а лишь оглашают некоторые их свойства, а в остальном предлагают полагаться на интуицию. Также придётся сделать и нам при изучении множеств.

2.1 Множества и операции над ними

Когда говорят, что задано множество A, под этим понимают, что A представляет собой совокупность объектов, игнорируя при этом какие либо отношения между этими объектами, в частности порядок; кроме того, один объект не может входить в

множество более одного раза. Конечное множество можно задать явно перечислив его элементы — для этого используют фигурные скобки:

$$\{1, 2, 3, 4, 5\}.$$

Из сказанного выше вытекает, что

$$\{1, 2, 3, 4, 5\} = \{5, 4, 3, 2, 1\} = \{1, 3, 2, 4, 5\} = \{1, 1, 2, 2, 2, 3, 4, 5\}.$$

Два множества ***равны*** друг другу, если их элементы совпадают. В последнем описании множества элементы повторяются: элементы разрешено повторять при перечислении, хотя каждый перечисленный элемент и входит в множество ровно один раз. Количество элементов конечного множества A называют ***мощностью*** A и обозначают через $|A|$: $|\{1, 2, 3\}| = 3$.

В случае описания бесконечных множеств, используют неявное перечисление: так множество $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$ состоит из (всех) ***натуральных чисел***. Мы считаем ноль натуральным числом (по этому вопросу среди математиков нет единого мнения), и чтобы не путать читателя обозначаем натуральные числа через $\mathbb{N}_0$; обозначим через $\mathbb{N}_1 = \{1, 2, 3, \ldots\}$ множество ***положительных целых чисел***. Множество ***целых чисел*** $\mathbb{Z}$ часто записывают одним из следующих способов:

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\} = \{0, 1, -1, 2, -2, \ldots\}$$

Запись «$a \in A$» означает, что объект a ***является элементом*** множества A, а запись «$a \notin A$» — отрицание этого условия:

$$3 \in \{1, 2, 3, 4, 5\}, \qquad 2 \notin \{1, 3, 5, \ldots\}.$$

Познакомимся с ещё одной формой записи множеств. Запись

$$A = \{x \mid \text{«условие на } x\text{»}\}$$

означает, что множество A состоит из всех элементов x, для которых выполняется условие. Так, запись $\{x \mid x = 2k + 1 \text{ для некоторого } k \in \mathbb{N}_0\}$ задаёт множество нечётных чисел. Также вместо символа $\mid$ используют двоеточие; эти записи равноправны, а выбор символа часто обусловлен красотой и читаемостью формулы. Определим с помощью такой записи множество ***рациональных чисел***:

$$\mathbb{Q} = \left\{ \frac{m}{n} \,:\, m \in \mathbb{Z},\, n \in \mathbb{N}_1 \right\}.$$

Также введём здесь обозначение для множества ***действительных чисел*** $\mathbb{R}$, аккуратное определение которых даётся в курсе математического анализа.

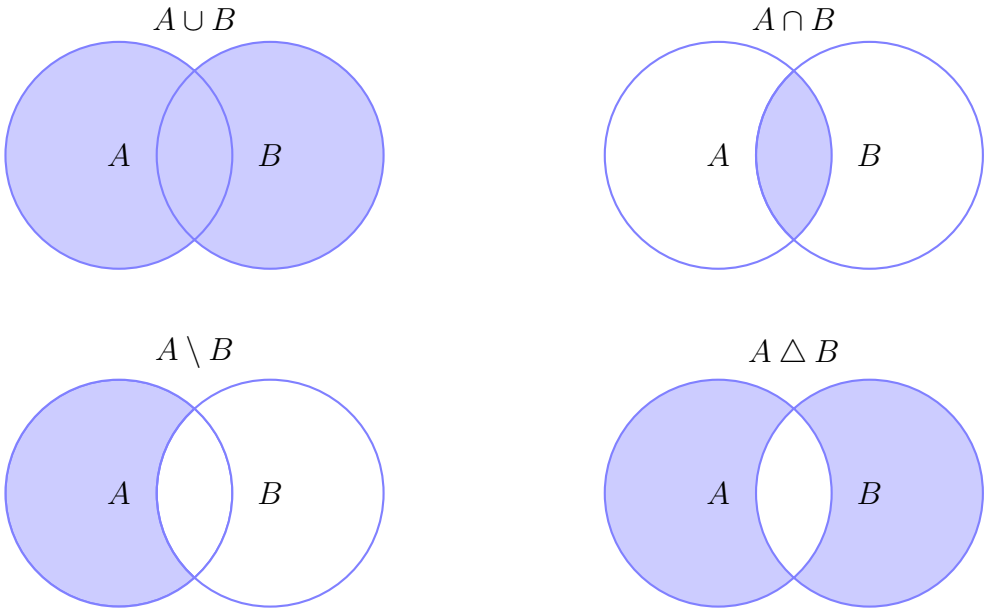


Рис. 2.1. Диаграммы Эйлера-Венна.

Определим теперь операции над множествами:

- объединение $A \cup B$ множеств A и B состоит из элементов, которые принадлежат хотя бы одному из множеств;
- пересечение $A \cap B$ состоит из элементов, которые принадлежат обоим множествам;
- разность $A \setminus B$ состоит из элементов, которые принадлежат A, но не принадлежат B;
- симметрическая разность $A \triangle B$ состоит из элементов, принадлежащих ровно одному из множеств.

Эти операции иллюстрируют с помощью диаграмм Эйлера-Венна (рис. 2.1).

Говорят, что множество C является ***подмножеством*** множества D, если каждый элемент множества C принадлежит D. Это обозначают $C \subseteq D$. Из картинок видно, что $A \setminus B \subseteq A$. Множество, в котором нет элементов называют ***пустым*** и обозначают $\varnothing$.

Упражнение 1. Убедитесь, что $(A \triangle B) \cap (A \cap B) = \varnothing$.

Упражнение 2. Убедитесь, что множества A и B равны, тогда и только тогда, когда $A \subseteq B$ и $B \subseteq A$.

Упражнение 3. Докажите, что для любых множеств A и B справедлива формула $(A \cup B) \setminus (A \triangle B) = A \cap B$.

2.2 Связь с логикой

При изучении алгебры логики, мы имели дело с высказываниями, которые либо истинны, либо ложны. Утверждение $A(x) =$ «x делится на 6» зависит от параметра x, а потому таковым не является. Утверждения, зависящие от параметров, называют ***предикатами*** и работать с ними можно точно так же, как и с обычными высказываниям (можно применять к ним всё те же логические связки). Часто удобно использовать предикаты, зависящие от нескольких параметров: например, $G(x, y) =$ «$x > y$». Число параметров называется ***арностью*** предиката, предикаты арности 1 или ***унарные*** предикаты соответствуют множествам: множеству A соответствует предикат $A(x)$, который истинен тогда и только тогда, когда $x \in A$.

Обратим внимание, что

$$A \cup B = \{x \mid (x \in A) \vee (x \in B)\};$$

$$A \cap B = \{x \mid (x \in A) \wedge (x \in B)\};$$

$$A \setminus B = \{x \mid (x \in A) \wedge \neg(x \in B)\};$$

$$A \triangle B = \{x \mid ((x \in A) \wedge \neg(x \in B)) \vee ((x \in B) \wedge \neg(x \in A))\}.$$

В описании множества $\{x \mid P(x)\}$ условие $P(x)$ — это и есть унарный предикат, задающий множество.

Эта связь объясняет законность рассуждений с картинками (диаграммами Эйлера-Венна). Допустим в диаграмму входят три множества A, B и C. Каждую область диаграммы можно задать вектором (a, b, c) с компонентами $\{0, 1\}$; значение 1 означает, что x принадлежит соответствующему множеству, а 0, что нет. Так, вектор $(1, 0, 1)$ означает, что $x \in A$, $x \notin B$ и $x \in C$. Сама диаграмма описывает множество D, и если область закрашена, то всякий x из этой области принадлежит D. Получаем, что диаграммы Эйлера-Венна просто иллюстрируют таблицы истинности.

Многие теоретико-множественные тождества следуют напрямую из тождеств алгебры-логики: тождество

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

равносильно следующему тождеству, в котором a означает «$x \in A$» и т.д.

$$(a \wedge b) \vee c = (a \vee c) \wedge (b \vee c).$$

Юнивёрсум и Дополнение

Вы скорее всего уже наслышаны о такой операции с множествами, как дополнение. Но вот каверзный вопрос[1]: принадлежит ли число $\sqrt{2}$ к дополнению множества чётных чисел? А вот этот стул? Слово «дополнение» буквально означает «до полного», поэтому для этой операции нужно сначала определить «полное множество» — множество из всех объектов, которые мы рассматриваем в рамках наших теоретико-множественных рассуждений. Такое множество называют ***юнивёрсум*** и обозначают U.

[1] Если о дополнении вы не слышали, то пропустите этот вопрос.

Множество $\bar{A} = U \setminus A$ называют ***дополнением*** к множеству A. Ясно, что $\bar{A}$ — это наименьшее множество, которое нужно добавить к A, чтобы получилось множество U.

Мы изучаем наивную теорию множеств, которая вообще говоря противоречива. Проблемы возникают, если рассматривать множества всех множеств (см. парадокс Рассела). Для того, чтобы обезопасить себя, достаточно зафиксировать юнивёрсум: если U — множество натуральных (целых неотрицательных) чисел $\mathbb{N}_0$, целых чисел $\mathbb{Z}$, рациональных чисел $\mathbb{Q}$ или вещественных чисел $\mathbb{R}$, то проблем не будет.

Приведём теперь законы теории множеств, связанные с дополнением и соответствующие им логические законы.

Законы де Моргана

С помощью диаграмм легко проверить, что $A \cap B = \overline{\bar{A} \cup \bar{B}}$, $A \cup B = \overline{\bar{A} \cap \bar{B}}$. Из связи с таблицами истинности получаем, что $a \wedge b = \neg(\bar{a} \vee \bar{b})$ и $a \vee b = \neg(\bar{a} \wedge \bar{b})$. Эти законы известны как законы де Моргана.

Однако, эти формулы можно обобщить:

$$A_1 \cap A_2 \cap \ldots \cap A_n \cap \ldots = \overline{\bar{A}_1 \cup \bar{A}_2 \cup \ldots \cup \bar{A}_n \cup \ldots} \tag{1}$$

Докажем обобщённую формулу, обозначим левую часть за X, а правую за Y. Если $x \in X$, то x принадлежит каждому множеству A_i, но тогда он не принадлежит ни одному дополнению $\bar{A}_i$, а значит и их объединению. Значит x принадлежит дополнению от объединения дополнений, т. е. Y. Мы доказали, что $X \subseteq Y$. Пусть теперь $y \in Y$, тогда $y \notin \overline{Y}$ и потому для каждого i выполняется $y \notin \bar{A}_i$. Но раз $y \notin \bar{A}_i$, то $y \in A_i$ (для каждого i), а потому $y \in X$. Отсюда $Y \subseteq X$; как и в первом случае включение справедливо в силу произвольности y. Итак, мы доказали, что $X = Y$, что и требовалось. Обратим внимание, что при доказательстве равенства двух множеств требуется доказывать включения в обе стороны! Бывают доказательства, в которых хитрым образом доказывается равенство множеств без доказательств включений по очереди, но это скорее редкость.

Двойственный закон Моргана

$$B_1 \cup B_2 \cup \ldots \cup B_n \cup \ldots = \overline{\bar{B}_1 \cap \bar{B}_2 \cap \ldots \cap \bar{B}_n \cap \ldots}$$

можно доказать аналогично, но можно и вывести из первого закона. Поскольку тождество (1) справедливо для произвольных множеств, заменим в нём A_i на $\bar{B}_i$, снимем двойное дополнение и возьмём дополнения от обеих частей равенств.

Кванторы

Возможно вы уже познакомились с кванторами в математическом анализе. Вернёмся к утверждениям, зависящим от параметра — предикатам. Часто интересно, истинен ли предикат A при любом x. Это утверждение записывают как

$$\forall x A(x),$$

а истинность при хотя бы одном x

$$\exists x A(x).$$

Значки $\forall$ и $\exists$ называют ***кванторами*** (всеобщности и существования соответственно).

Кванторы можно интерпретировать как (возможно) бесконечные конъюкции и дизъюнкции элементарных высказываний $A(x)$. Поскольку операции конъюнкция и дизъюнкция коммутативны и ассоциативны, порядок их выполнения не важен. Это приводит к следующему сокращению в формулах:

$$A(1) \wedge A(2) \wedge \dots A(n) = \bigwedge_{i=1}^{n} A(i) = \bigwedge_{i \in \{1,\dots,n\}} A(i).$$

В случае когда порядок операндов важен (например, в произведении матриц), вторая запись интерпретируется как первая, а третья запись вообще говоря некорректна.

В случае произвольной формулы в кванторах, подразумевается что каждая переменная принимает значение из определённого множества, юнивёрсума U. В логике удобно считать, что все переменные принимают значения из единственного множества, это легко реализовать технически.

Итак, формально формула в кванторах интерпретируется так:

$$\forall x A(x) = \bigwedge_{x \in U} A(x).$$

К конъюнкции (не обязательно конечной) применим закон Моргана, отсюда получаем, что

$$\neg \forall x A(x) = \overline{\bigwedge_{x \in U} A(x)} = \bigvee_{x \in U} \neg A(x) = \exists x \neg A(x).$$

Эквивалентность тождеств алгебры логики и алгебры множеств

Тождества алгебры логики переходят в тождества алгебры множеств при замене булевых переменных множествами, а операций алгебры логики на соответсвующие операции алгебры множеств. Переход справедлив и в обратную сторону. Поэтому все преобразования, которые мы изучили работая с алгеброй логики, имеют прямой аналог в алгебре множеств.

Приведём набросок доказательства эквивалентности тождеств в этих алгебрах. Возьмём формулу алгебры множеств и заменим в ней множества A_i на предикаты $A_i(x)$ а операции, на соответствующие логические операции. Добавив перед обеими частями квантор всеобщности по x и заменив равенство на эквивалентность получим выражение вида

$$\forall x((\text{левая часть формулы}) \leftrightarrow (\text{правая часть формулы})).$$

Ясно, что это утверждение истинно для любых предикатов тогда и только тогда, когда изначальная формула алгебры множеств справедлива для любого набора множеств. Также ясно, что это условие выполняется, если заменив теперь предикаты на булевы переменные ($A_i(x)$ на a_i) и убрав квантор по x, мы получим

тождество в алгебре логики. Если же в результате замены мы получили не тождество, найдётся такой набор переменных, при котором высказывание ложно. Если в этом наборе $a_i = 1$, положим $A_i = \{1\}$, если же $a_i = 0$, положим $A_i = \varnothing$. Выполнив обратную замену от алгебры логики к утверждению в предикатах и к формуле алгебры множеств, получим, что утверждение в предикатах ложно, а формула алгебры множеств не выполняется.

Рассуждения в обратную сторону аналогичны.

Приведём пример построения эквивалентных тождеств алгебры множеств и алгеброй логики с промежуточным шагом формулы с предикатами на примере закона де Моргана.

Пример 3.

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

$$\forall x(\neg(A(x) \vee B(x)) \leftrightarrow (\neg A(x) \wedge \neg B(x)))$$

$$\neg(a \vee b) \leftrightarrow \neg a \wedge \neg b$$

Импликация и множества

На первый взгляд, определение импликации выглядит странно. Почему математики решили, что $0 \to 1 = 1$?! Ответ кроется в связи с теорией множеств. Утверждение «(x делится на 6) $\to$ (x делится на 2)$\vee$(x делится на 3)» будет считаться теоремой, если оно истинно при всех x:

$$\forall x\ D_6(x) \to (D_2(x) \vee D_3(x)).$$

Но утверждение $D_6(x)$ ложно, например, при $x = 4$, а утверждение $D_2(x)$ истинно.

Переведём выссказывание $\forall x\ A(x) \to B(x)$ на язык множеств. Если посылка импликации истинна при некотором x, то при этом же x истинно и заключение (иначе, импликация ложна). Если же посылка ложна, то какого бы ни было заключение, импликация истинна. Значит каждый элемент x из множества A принадлежит также и множеству B, отсюда получаем, что множество A — подмножество множества B. Ясно, что $D_6 \subseteq D_2 \cup D_3$.

Контрапозиция

Логический закон контрапозиции $A \to B = \neg B \to \neg A$ при переводе на язык множеств гласит, что $A \subseteq B \iff \bar{B} \subseteq \bar{A}$. Его часто используют на практике при доказательстве теорем и решении задач: когда нужно доказать следствие $A \to B$, часто вместо него доказывают $\neg B \to \neg A$.

Лекция 3

Математические определения, утверждения и доказательства

План:

1. Определение, утверждение, теорема, критерий. Запись с помощью формулы первого порядка (неформально).
2. Методы доказательств: контрапозиция, индукция, от противного, конструктивные (примеры и контрпримеры), неконструктивные.
3. Границы применимости: парадокс Рассела.

Литература: [4], [5], [6], [1]

Изучив основы логики и теории множеств мы можем содержательно поговорить о доказательствах. Наш разговор не будет строгим; строгому изложению этого материала отведено место на втором курсе, но изучать доказательства и что-то доказывать при решении задач, нужно уже сейчас.

3.1 Определения

Определения описывают объекты и понятия. Если определение записано логической формулой, то оно имеет вид предиката $D(x)$, который истиннен тогда и только тогда, когда x, удовлетворяет определению.

Пример 4. Множеству $D = \{x \mid x^2 + 2x + 1 = 0\}$ соответствует предикат $D(x)$, который определяет корни многочлена $x^2 + 2x + 1$, т. е. -1.

Пример 5. Формула

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N}_1 : \forall n \geqslant N \; |x_n - a| < \varepsilon$$

Как хорошо известно читателю, определяет предел числовой последовательности. Формально это предикат $D(a, \{x_n\})$, который зависит как от числа a, так и от последовательности $\{x_n\}$. Параметры, от которых зависит истинность формулы, не стоят под кванторами.

Определения, данные словами ничуть не хуже определений, данных формулами. На первом курсе последние встречаются чаще, чтобы научить студентов изложению в кванторах. Так, определение предела можно переформулировать словами: «число a — предел последовательности $\{x_n\}$, если любая окрестность числа a содержит все элементы последовательности, начиная с некоторого номера».

3.2 Математические утверждения

Математические утверждения — это утверждения, которые либо, истинны либо ложны. В отличие от определений, они не зависят от параметров. Если вы встретили утверждение вида «если последовательность x_n сходится, то она ограничена», то в силу вступает математическое соглашение о том, что в случае отсутствия в утверждении квантора по параметру, нужно поставить квантор всеобщности.

Среди математических утверждений выделяют ***теоремы*** — истинные утверждения. Как правило, теоремами называют значимые математические утверждения. Вспомогательные истинные математических утверждения называют ***леммами***, ***предложениями*** и просто ***утверждениями***.

Истинное утверждение называют ***критерием***, если оно имеет вид

$$\forall x\ (A(x) \leftrightarrow B(x)).$$

Критерии устанавливают необходимое и достаточное условие $B(x)$ для выполнения условия $A(x)$ или, что то же самое, устанавливает эквивалентность определений A и B. Например, в математическом анализе критерий Коши устанавливает эквивалентность сходящихся и фундаментальных последовательностей.

Рассмотрим утверждения вида

$$\forall x(A(x) \to B(x)) \tag{1}$$

Условие $B(x)$ является ***необходимым*** для выполнения $A(x)$, а условие $A(x)$ является ***достаточным*** для выполнения $B(x)$. Условие $A(x)$ считатеся более ***сильным***, чем $B(x)$, а $B(x)$ считается более ***слабым***, чем $A(x)$.

Смысл этих определений вытекает из определения импликации, напомним что уравнение (1) на языке множеств означает, что $A \subseteq B$, где A и B множества, соответствующие предикатам ($A = \{x \mid A(x)\}$). Отсюда вытекает, что условие $B(x)$ выполняется всегда, когда выполняется $A(x)$, отсюда $A(x)$ — достаточное условие; если условие $A(x)$ выполняется, то всегда выполняется и $B(x)$, отсюда $B(x)$ — необходимое условие. Условие $A(x)$ считается сильнее условия $B(x)$, потому что все условия, которые следуют из выполнения $B(x)$, также следуют и из выполнения $A(x)$:

$$\{C(x) \mid \forall x(A(x) \to C(x))\} \supseteq \{C(x) \mid \forall x(B(x) \to C(x))\}.$$

Отметим также, что в случае теорем вида $\forall x(A(x) \to C(x))$ и $\forall x(B(x) \to C(x))$ вторая теорема считается сильнее первой, потому что в ней условие $C(x)$ следует из более слабого условия $B(x)$.

3.3 Доказательства

Доказательство — это логическое рассуждение, которое убеждает в верности математического утверждения любого непредвзятого слушателя (читателя). У доказательств есть формальное определение в математической логике, но оно требует введение формальных систем и фактически такие доказательства непроверяемы человеком. Математики любят пользоваться приведённым описанием доказательства, но в утилитарном смысле оно слабо годится. Откуда первокурснику знать, убедят ли его аргументы академика? Поэтому помимо философского описания, мы дадим ещё и утилитарное, но для этого нам потребуется сначала описать логический вывод.

Логический вывод

Представьте, что известна истинность утверждений A и $A \to B$. Из этого можно сразу заключить истинность утверждения B, ведь если B ложно, а A истинно, то импликация $A \to B$ ложна. В формальной логике у этого правила есть своё имя (Modus Ponens), а у правил вывода есть специальная запись:

$$\frac{A, \quad A \to B}{B} \tag{М.Р.}$$

Запись интерпретируется так: если доказано то, что выше черты, то доказано и то, что ниже черты. По аналогии с импликацией, то что выше черты называют посылкой, а то что ниже — заключением. Мы не будем уделять внимание разным правилам вывода и акцентировать внимание на этой записи — мы привели их здесь, чтобы описать общие идеи.

Первая состоит в том, что если известна истинность какого-то сложного (составного) логического высказывания, то используя преобразования формул или логические рассуждения можно доказать истинность или ложность частей этого выссказывания, вплоть до элементарных высказываний (логических переменных). Например, пусть известно, что следующее высказывание истинно

$$(\neg A) \wedge (A \vee B) \tag{2}$$

Отсюда сразу следует, что истинны операнды конъюнкции: $\neg A$ и $A \vee B$. Из истинности первого операнда следует, что $A = 0$, а из этого факта и истинности второго операнда следует $B = 1$. В процессе решения задач и доказательства теорем, не обязательно известна истинность сложного выссказывания, как правило известна истинность нескольких фактов, например $\neg A$ и $A \vee B$, из которых можно составить сложное высказывание (2) и с помощью логических преобразований получить

результаты, которые мы получили, но в то же время можно и не составлять формулу (2), а использовть логическое рассуждение. Поэтому в логике, полученное нами правило вывода записали бы так:

$$\frac{\neg A, \quad A \vee B}{B}.$$

Отметим, что это правило вывода, как и многие другие, сводится к Modus Ponens: $A \vee B = \neg A \to B$.

Формально запись

$$\frac{A_1, \quad A_2, \quad \ldots \quad A_n}{B}$$

означает, что

$$A_1 \wedge A_2 \wedge \ldots \wedge A_n \to B. \tag{3}$$

Если известно, что все утверждения A_i истинны, и истинно утверждение (3), то эти факты в совокупности влекут (доказывают) истинность утверждения B.

В этой части лекции мы использовали форму (формулы записи логического вывода), чтобы раскрыть суть. При доказательстве на каждом шаге у вас есть множество (список) истинных утверждений. Часть из них берётся из условия доказываемого утверждения, часть — факты из курса, а какую-то часть вы уже доказали. Используя логические рассуждения (преобразования, логические тождества, таблицы истинности) вы можете доказать истинность новых рассуждений, и расширить тем самым этот список. Расширять список нужно до тех пор, пока в нём не окажется доказываемое утверждение. Описанный процесс и является нашим ***требованием к доказательствам***. При работе с доказательствами (решении задач) нужно следовать сути этого метода, а не форме записи логического вывода, которая вторична.

Приведём пример.

Пример 6. Алису, Вениамина и Сергея вызвали к директору, потому что кто-то из них на перемене разбил окно. Алиса сказала, что ни она, ни Вениамин окно не разбивали. Вениамин сказал, что Алиса не разбивала окно, а это сделал Сергей, а Сергей сказал, что он не разбивал окно и окно разбила Алиса.

Директору известно, что ровно один школьник сказал правду, другой солгал в каждом из утверждений, а третий дал одно истинное, а другое ложное утверждение. Кто же разбил окно?

Решение. Обозначим через A, B, C высказывания «Алиса разбила окно», «Вениамин разбил окно», «Сергей разбил окно». Точно известно, что истинно высказывание

$$A \oplus B \oplus C.$$

Среди следующих высказываний истинно ровно одно, ещё в одном истинен ровно один конъюнкт, а в оставшемся ложны оба конъюнкта:

$$\neg A \wedge \neg B, \qquad \neg A \wedge C, \qquad \neg C \wedge A.$$

Предположим, что Алиса сказала правду. Тогда истины высказывания $\neg A$ и $\neg B$. Получаем отсюда, что окно разбил Сергей:

$$\frac{\neg A, \quad \neg B, \quad A \oplus B \oplus C}{C}.$$

Но это невозможно, потому что тогда Вениамин тоже сказал правду:

$$\frac{\neg A, \quad C}{\neg A \wedge C}.$$

Предположив, что правду сказал Вениамин, также получим, что окно разбил Сергей, и Алиса тоже сказала правду, что невозможно.

Получается, что правду сказал Сергей и окно разбила Алиса. На этом решение можно было бы закончить, при условии доверия к составителю задачи. Если быть формальными до конца, то нужно проверить оставшиеся условия. Ясно, что Алиса соврала наполовину (ровно одно из её высказываний истинно), а Вениамин соврал в каждом из утверждений. □

Заметим, что если записать условие примера с помощью формулы, то она получится очень длинной, и придётся мучиться с её упрощением. При доказательствах следует использовать логику по сути, а не пытаться всё формализовать излишне, если это путает дело.

Повторим, что наши требования к доказательствам (решениям) относятся к сути, а не к форме. Текст на естественном языке, удовлетворяющий им, ничуть не хуже (а часто лучше), чем набор формул с шагами вывода. Но при написании текста нужно понимать, какие утверждения в нём делаются, и какая между ними логическая связь; полезно помогать себе и читателю доказательства, явно выделяя вспомогательные утверждения.

Когда много доказательств используют одну и ту же структуру (опираются на одинаковую тавтологию), выделяют метод доказательства. Мы переходим к перечислению различных методов доказательств и примерам их применения.

Контрапозиция

Закон контрапозиции основан на тавтологии

$$(A \to B) \leftrightarrow (\neg B \to \neg A).$$

Он гласит, что утверждение $A \to B$ равносильно (контрапозитивному) утверждению $\neg B \to \neg A$, поэтому если требуется доказать первое, вместо него достаточно доказать последнее.

Смысл закона контрапозиции становится ясным при переходе на язык множеств (как и его справедливость): $A \subseteq B$ тогда и только тогда, когда $\overline{B} \subseteq \overline{A}$.

Приведём пример его использования.

Утверждение 1. *Если число r иррационально, то и число $\sqrt{r}$ иррационально.*

Доказательство. Воспользовавшись контрапозицией получим равносильное утверждение:

«Если число $\sqrt{r}$ рационально, то число r рационально.»

Это утверждение доказать нетрудно: если число $\sqrt{r}$ рационально, то $\sqrt{r} = \frac{m}{n}$, отсюда $r = \frac{m^2}{n^2}$ и получаем, что число r рационально по определению. □

Индукция

Отдельную сложность у студентов (увы, не только первокурсников) вызывают доказательства по индукции.

Доказательство по индукции возможно только тогда, когда доказываемое утверждение зависит от натурального параметра. То есть доказывается утверждение

$$\forall n \in \mathbb{N}_0 : A(n).$$

С помощью правил вывода схему доказательства по индукции можно описать так:

$$\frac{A(0), \quad \forall n \, \big(A(n) \to A(n+1)\big)}{\forall n \, A(n)}.$$

Первая посылка называется ***базой***, а вторая — ***шагом*** индукции или ***переходом***.

Заметим, что в случае утверждений, записанных формулой с кванторами (такие формулы называют формулы первого порядка), проверить истинность утверждения не всегда просто. Описанную с помощью вывода тавтологию

$$A(0) \wedge \forall n \, \big(A(n) \to A(n+1)\big) \to \forall n \, A(n)$$

либо причисляют к аксиомам (утверждениям, истинным по определению), либо выводят из других аксиом. Поэтому простого и достаточно строгого обоснования метода математической индукции дать не получится. Неформальное обоснование метода фактически опирается на сам метод: ясно, что если утверждения $A(0)$ и $A(0) \to A(1)$ истинны, то по М.Р. получаем, что и утверждение $A(1)$ истинно, а далее по индукции, т. е. при фиксированном n и истинных $A(n)$ и $A(n) \to A(n+1)$ получаем истинность $A(n+1)$.

Пример 7. Для каждого целого $n > 0$ справедливо

$$1 + 3 + 5 + \cdots + (2n-1) = n^2.$$

Доказательство. Обозначим доказываемое равенство через $A(n)$ и докажем его по индукции. Оговоримся сразу, что без ограничения общности можно считать, что база начинается не обязательно с нуля, как в этом примере, потому что вместо доказательства справедливости утверждения $A(n)$ с единицы можно было бы доказывать утверждение $B(m) = A(m+1)$ с нуля.

База: при $n = 1$ утверждение $A(1)$ утверждает равенство $1 = 1^2$, которое справедливо. Докажем переход; утверждение $A(n+1)$ гласит

$$1 + 3 + 5 + \cdots + (2n-1) + (2n+1) = (n+1)^2.$$

Считая утверждение $A(n)$ верным, получаем цепочку равенств

$$\underbrace{1 + 3 + 5 + \cdots + (2n-1)}_{=n^2 \text{ по утверждению } A(n)} + (2n+1) = n^2 + 2n + 1 = (n+1)^2,$$

которая доказывает утверждение $A(n+1)$. □

Более подробному изложению индукции и связанных с нею проблемам посвящена первая глава книги [1].

От противного

Мы полагаем, что если утверждение B истинно, то оно не может быть одновременно ложным. Если предположить, что утверждение A ложно и с помощью него доказать, что ложно утверждение B, то есть доказать истинность $\neg A \to \neg B$, то в случае, если утверждение B истинно, утверждение A не может быть ложным — иначе бы мы получили истинность B и $\neg B$. Отсюда вытекает способ доказательства от противного, который можно описать как

$$\frac{\neg A \to \neg B, \quad B}{A}.$$

Классический пример такого доказательства — иррациональность числа $\sqrt{2}$.
Доказательство. Доказательство от противного. Положим, что $\sqrt{2} = \frac{m}{n}$, где $\frac{m}{n}$ — несократимая дробь, $m \in \mathbb{Z}$, $n \in \mathbb{N}_1$. Тогда $m^2 = 2n^2$, отсюда m^2 делится на 2, и m делится на 2, значит m^2 делится на 4, и отсюда n^2 делится на 2 и n делится на 2. Но тогда и m делится на 2 и n делится на 2, а значит дробь $\frac{m}{n}$ сократима, пришли к противоречию. □

Примеры и контрпримеры

В случае если утверждение имеет вид $\exists x : A(x)$, его можно доказать, приведя ***пример*** (и доказав справедливость этого примера). Рассмотрим утверждение:

$$\exists n \in \mathbb{N}_0 : \sqrt{n} \in \mathbb{R} \setminus \mathbb{Q},$$

то есть существует натуральное число n, корень из которого — иррациональное число. Это утверждение очевидно верно, и для его доказательства достаточно предъявить число $n = 2$ и доказать иррациональность числа $\sqrt{2}$.

Рассмотрим теперь утверждение

$$\forall n \in \mathbb{N}_0 : \sqrt{n} \in \mathbb{R} \setminus \mathbb{Q}.$$

Это утверждение, очевидно, неверно: достаточно взять $n = 4$ и показать, что $\sqrt{4} = 2 \in \mathbb{Q}$. Для опровержения утверждения с квантором всеобщности $\forall x : A(x)$ достаточно привести ***контрпример***, т. е. пример x, для которого $A(x) = 0$.

Заметим, что для доказательства утверждений вида $\forall x : A(x)$ одного примера не достаточно. Даже если утверждение $A(x)$ верно при каком-то x или очень многих x, даже если их бесконечно много, отсюда ещё не вытекает, что утверждение $A(x)$ верно при всех x. Если все x не проверены, то возможно среди не рассмотренных есть контрпример. Но как проверить бесконечно много x? Вот несколько рецептов. Провести доказательство утверждения $A(x)$, которое не зависит от выбора x. Если x пробегает счётное множество значений (т. е. $\mathbb{N}_0$ или другое множество, элементы которого можно занумеровать), то можно воспользоваться индукцией. Воспользоваться методом доказательства от противного: предположить $\exists x : \neg A(x)$ и прийти к противоречию.

Неконструктивные доказательства

Утверждение вида $\exists x : A(x)$ не обязательно доказывать приводя пример, хотя это очень желательно, если таковой имеется — наличие примера или контрпримера лучше всего убеждает в справедливости утверждения. Бывает так, что само утверждение доказать проще, чем найти пример и мы приведём здесь такое доказательство.

Утверждение 2. *Существуют иррациональные числа a и b, такие что число a^b рационально.*

Доказательство. Положим, что $a = b = \sqrt{2}$. Если число $(\sqrt{2})^{\sqrt{2}}$ рационально, то утверждение доказано. Если нет, то возьмём $a = (\sqrt{2})^{\sqrt{2}}$, а $b = \sqrt{2}$:

$$\left(\sqrt{2}^{\sqrt{2}}\right)^{\sqrt{2}} = \left(\sqrt{2}\right)^{\left(\sqrt{2}\times\sqrt{2}\right)} = \left(\sqrt{2}\right)^{2} = 2.$$

То есть, либо подходит одна пара чисел, либо другая, а какая из — мы не знаем.

Обычно неконструктивные доказательства приводят в некоторое замешательство, особенно при первом знакомстве. Разберёмся со структурой доказательства, формализовав рассуждения.

Само утверждение имеет вид $\exists a, b : A(a, b)$. Мы предположили сначала, что справедливо утверждение $A(\sqrt{2}, \sqrt{2})$, если же оно неверно, то мы доказали, что отсюда вытекает утверждение $A((\sqrt{2})^{\sqrt{2}}, \sqrt{2})$. То есть мы доказали утверждение:

$$\neg A(\sqrt{2}, \sqrt{2}) \to A((\sqrt{2})^{\sqrt{2}}, \sqrt{2}).$$

Перейдя от импликации к дизъюнкции, получаем

$$A(\sqrt{2}, \sqrt{2}) \vee A((\sqrt{2})^{\sqrt{2}}, \sqrt{2}).$$

Доказанная дизъюнкция очевидно влечёт доказываемое утверждение $\exists a, b : A(a, b)$. □

3.4 Границы применимости

Хорошо известно, что у физических теорий есть границы применимости: при больших скоростях законы Ньютоновской механики становятся неприменимы, и требуется использовать теорию относительности, а в микромире нужна квантовая механика. Удивительно, но границы применимости есть и у логики, а точнее у теории множеств. За две последние лекции вы должны были убедиться, что эти области тесно связаны.

Мы изучили наивную теорию множеств (не заботились об аксиоматических определениях), которая вообще говоря неверна, и в качестве иллюстрации этого часто приводят парадокс Рассела, известный также как парадокс брадобрея, который звучит так.

Есть деревня, в которой бреется каждый мужчина. При этом либо каждый мужчина бреется сам, либо его бреет единственный в деревне брадобрей (парикмахер). Таким образом, брадобрей бреет тех и только тех мужчин, которые не

бреются сами. Парадокс состоит в том, что бреет ли себя брадобрей? Если да, то он бреется сам, значит брадобрей, то есть он же, брить себя не должен. Если он не бреется сам, то его обязан брить брадобрей.

На язык теории множеств этот парадокс переводится так. «Содержит ли себя множество, которое содержит в качестве элемента каждое множество, которое не содержит себя (в качестве элемента)?» И в случае ответа «да», и в случае ответа «нет» приходим к противоречию.

Как же нам вести доказательства, если мы используем наивную теорию множеств, которая вообще говоря противоречива? Хорошие новости состоят в том, что с момента обнаружения проблем, математики занимались формализацией теории множеств, разработали формальную логику, и мы можем использовать плоды их трудов для наших скромных целей. Как видно, проблемы парадокса начинаются, когда мы рассматриваем неограниченный юнивёрсум (множество всех множеств в принципе). В наших рассуждениях мы всюду будем использовать ограниченные и довольно скромные (по мощности) юнивёрсумы. Поэтому подход к доказательствам, принятый в нашем курсе безопасен для наших целей.

Лекция 4

Графы I. Простые неориентированные графы

План:

1. Определение неориентированных графов
2. Степень вершины. Сумма степеней вершин — удвоенное количество рёбер.
    - Число людей, сделавших нечётное число рукопожатий, чётно.
3. Теоретико множественные операции с графами. Определение подграфа
4. Определение путей и циклов (через подграфы)
5. Связные графы и компоненты связности (через подграфы)

Литература: [7], [8], [1]

Мы переходим сейчас к изучению графов по двум причинам. Во-первых, графы иллюстрируют как можно легко ввести новые определения пользуясь аппаратом теории множеств, а во вторых графы — прекрасный полигон для упражнений в доказательствах.

Неформально определение графа легко объяснить с помощью картинок: нарисуем точки, соединим их линиями и получим тем самым граф. Сразу нужно оговориться, что пересечения линий смысла не несут и то, что мы соединяем линиями только разные точки, и каждую пару точек соединяем линией не более одного раза.

Понятие графа легко формализовать с помощью теории множеств. Точки образуют множество ***вершин*** V (произвольное множество любой природы), а линии формализуются как его двухэлементные подмножества $\{u, v\}$, которые образуют множество ***рёбер*** E. Поскольку каждое ребро состоит ровно из двух

элементов, то нет рёбер из вершины в себя, они бы имели вид $\{v\}$, такие рёбра называются ***петлями***; поскольку множество содержит каждый элемент не более одного раза, то двух рёбер между одной и той же парой вершин быть не может, такие рёбра называют ***кратными*** или ***параллельными***.

Введём вспомогательное обозначение. Обозначим через $\binom{A}{2}$ все двухэлементные подмножества множества A, т. е.

$$\binom{A}{2} = \{\{a, b\} \mid a, b \in A,\ a \neq b\}.$$

Формализуем определение графа. (Простой, неориентированный) ***граф*** G — состоит из множества ***вершин*** V и ***рёбер*** $E \subseteq \binom{V}{2}$; формально, граф — это упорядоченная пара $G = (V, E)$. Будем ссылаться на множество вершин и множество рёбер графа G через $V(G)$ и $E(G)$, даже если при определении графа G множества вершин и множество рёбер были обозначены другими буквами. Если множества вершин V и рёбер E заданы, то граф на них обозначают $G(V, E)$, это обозначение используют также чтобы быстро ввести множество вершин и множество рёбер графа: запись $H(W, I)$ означает, что $W = V(H), I = E(H)$.

Если не оговорено противного, то под графом мы понимаем простой неориентированный граф. ***Простой*** означает, что в графе нет петель и кратных рёбер, а ***неориентированный***, что рёбра графа не имеют направлений. В случае если заменить в графе линии на стрелки, т. е. ребро — упорядоченная пара вершин, то получится ***ориентированный граф***. Заметим, что множество вершин графа может вообще говоря быть бесконечным или пустым, но если не оговорено противного, то мы считаем, что множество вершин конечно и непусто.

4.1 Вершины, рёбра, степени вершин

Зафиксируем граф $G(V, E)$. Вершины u и v называются ***смежными*** или ***соседями***, если они образуют ребро: $\{u, v\} \in E$. Рёбра e и f называются ***смежными***, если они имеют общую вершину: $e \cap f \neq \varnothing$. Вершина u ***инцидента*** ребру e, если $u \in e$. Вершины u и v, инцидентные ребру e, называются его ***концами***; говорят, что e ***соединяет*** u и v. Рёбра часто записывают сокращённо: uv вместо $\{u, v\}$.

Степенью вершины v называется число смежных с v рёбер и обозначается $d(v)$.

Теорема 1. $\sum\limits_{u \in V} d(u) = 2|E|$.

Доказательство. В левой сумме ребро $\{u, v\}$ было подсчитано два раза: один раз в слагаемом $d(u)$, а в другой раз в слагаемом $d(v)$. □

Из теоремы сразу вытекает следующее следствие:

Следствие 1. *В любом графе число вершин нечётной степени чётно.*

Доказательство. В правой части равенства (из условия теоремы) стоит чётное число. Вычтем из обеих частей равенства все чётные степени вершин и получим, что в левой части осталась сумма нечётных степеней вершин, а в правой — чётное число. □

Это тривиальное следствие позволяет доказывать следующие странные утверждения, такие как «число людей в этой аудитории, пожавших с утра руки нечётному числу людей (в этой аудитории), чётно».

4.2 Базовые графы

Приведём примеры графов, которые встречаются в теории графов так часто, что получили собственные имена.

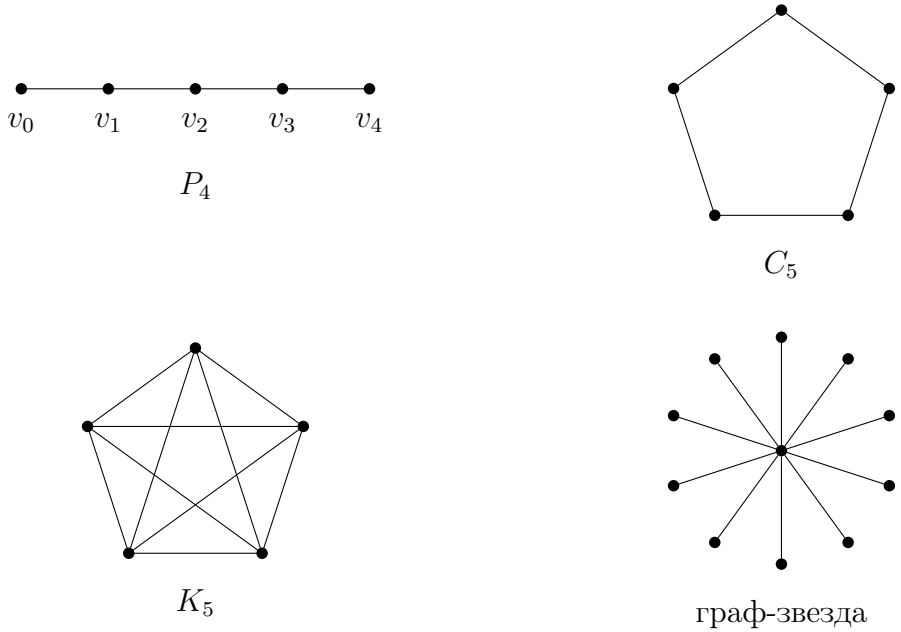


Рис. 4.1. Базовые графы

Граф-путь P_n, $n \geqslant 0$ состоит из вершин $\{v_0, v_1, \ldots, v_n\}$ и рёбер $\{v_i, v_{i+1}\}$. ***Длина пути*** — это число рёбер в пути, которых в графе P_n ровно n (поэтому нумерация вершин графа начинается с нуля); вершина v_0 называется ***началом пути***, а вершина v_n — ***концом пути***. Заметим, что граф P_0 состоит из одной вершины и не имеет рёбер и полноправно считается путём длины 0. Мы называем графом-путём любой граф, представимый в описанном виде (природа множества вершин нас не интересует), это относится и к остальным графам с рис. 4.1.

Граф-цикл C_n, $n \geqslant 3$ состоит из вершин $v_1, \ldots, v_n$ и рёбер $\{v_i, v_{i+1}\}$, а также $\{v_n, v_1\}$. Как и в случае пути, длина цикла — это количество рёбер в цикле.

Полный граф $K_n(V, E)$, $n \geqslant 1$ состоит из n вершин и имеет всевозможные рёбра: $E = \binom{V}{2}$.

Граф-зведа состоит из выделенной вершины, соединённой рёбрами со всеми остальными вершинами (больше рёбер в этом графе нет).

Пустой граф не содержит ни вершин, ни рёбер: $V = E = \varnothing$ и обозначается как $\varnothing$. Во всех утверждениях о графах мы полагаем, что граф G не пуст; пустой граф удобно использовать в формулах, например, чтобы кратко сказать, что графы G и H не имеют общих вершин и рёбер.

Обратим внимание, что если граф подпадает под одно из определений выше, то это вовсе не означает, что он не подпадает под другое. Так граф-путь P_1 является одновременно графом-звездой, а граф-цикл C_3, ***треугольник***, является также полным графом K_3.

4.3 Теоретико-множественные операции с графами. Подграфы

Операции из теории множеств переносятся на графы естественным образом. Определим операции объединение, пересечение и дополнение графов, с помощью графов $G(V, E)$ и $H(W, I)$:

$$G \cup H = (V \cup W,\, E \cup I), \quad G \cap H = (V \cap W,\, E \cap I), \quad \overline{G} = \left(V,\, \binom{V}{2} \setminus E\right).$$

Операции объединения и пересечения определены естественным образом, а дополнение графа G — это минимальный граф $\overline{G}$ на том же множестве вершин, который при объединении с G даёт полный граф. Под минимальностью понимается, что из $\overline{G}$ нельзя удалить ребро и получить граф, который в объединении с G даст полный граф, это условие равносильно $E(G) \cap E(\overline{G}) = \varnothing$.

На рисунке 4.2 пример теоретико-множественных операций со знакомыми графами.

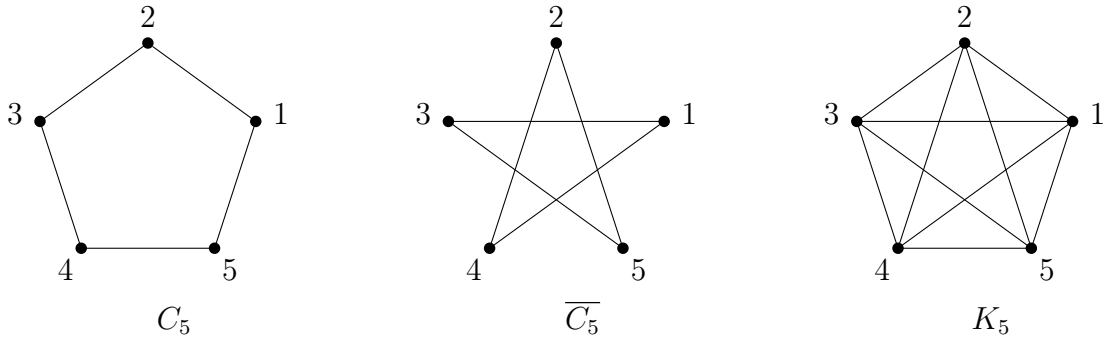


Рис. 4.2. Объединение графа и его дополнения даёт полный граф

Объединив граф C_5 с его дополнением получается граф K_5. При работе с теоретико-множественными операциями важно явно описать множество вершин каждого графа. Так, если бы у второго графа на картинке вместо вершин $\{1, 2, \ldots, 5\}$ были вершины $\{a, b, \ldots, e\}$, то в результате объединения получился бы другой граф. Также напомним, что обозначения вида C_n обозначают, что какой-то граф является графом-циклом. Так, граф $\overline{C_5}$ на картинке является графом C_5 — проверьте это перенумеровав его вершины.

Мы не определили на графах теоретико-множественную операцию разности (и соответственно симметрическую разность), потому что в зависимости от нужд требуется либо удалять из одного графа рёбра другого вместе с вершинами, либо вершины нужно оставить; чтобы не было путаницы мы будем явно указывать какую разность используем, оперируя множествами вершин и рёбер избегая сокращений.

Подграфы

Граф $H(W, I)$ называется ***подграфом*** графа $G(V, E)$, если $W \subseteq V$ и $I \subseteq E$. Другими словами, граф H получается из графа G удалением рёбер и вершин (вместе со смежными рёбрами). Это обозначают $H \subseteq G$. Мы используем определение подграфа, также известное как ***рёберный подграф***.

Формально, граф G является своим подграфом; чтобы подчеркнуть, что подграф H не совпадает с G используют обозначение $H \subsetneq G$; в этом случае подграф H называют ***собственным*** подграфом графа G.

Среди подграфов есть важный частный случай. Пусть $U \subseteq V$; подграф H графа G, состоящий из вершин U и содержащий все рёбра, которые есть в G называется ***индуцированным*** (множеством U); формально $H = \left(U, E \cap \binom{U}{2}\right)$. Для подграфа графа G индуцированного множеством U используют обозначение $G[U]$. По-умолчанию под подграфами мы подразумеваем рёберные подграфы.

С помощью понятия подграфа и базовых графов вводится ряд важных определений. Подграф H графа G называется

- ***путём*** из вершины u в вершину v, если H — это граф-путь P_n с началом в u и концом в v;
- ***циклом***, если H — это граф-цикл C_n;
- ***кликой***, если H — это полный граф K_n.

Множество $U \subseteq V(G)$ называется ***независимым***, если в индуцированном U подграфе нет рёбер, то есть $G[U] = \overline{K_n}$. Другими словами, никакие две вершины множества U не соединены ребром в графе G (это эквивалентное определение).

Перед следующим определением напомним, что формально объект обладает свойством $\mathcal{H}$, если он принадлежит множеству, определяемому этим свойством, которое мы обозначаем также, то есть $\mathcal{H} = \{x \mid \mathcal{H}(x)\}$; свойство формализовано через предикат $\mathcal{H}(x)$.

Пусть $\mathcal{H}$ — это некоторое свойство графов; обозначим через $\mathcal{H}_G$ все подграфы графа G, обладающие свойством $\mathcal{H}$. Подграф $H \subseteq G, H \in \mathcal{H}_G$ называется ***максимальным*** среди подграфов со свойством $\mathcal{H}$, если не существует подграфа $H' \in \mathcal{H}_G$, такого что $H \subsetneq H'$ и $H' \subseteq G$.

Пример 8. Рассмотрим граф K_5 на рис. 4.2. Рассмотрим свойство «быть циклом». Любой подграф графа K_5, являющийся циклом будет максимальным циклом, потому что добавлением вершин и рёбер в цикл, нельзя получить другой цикл; таким образом графы C_5 и $\overline{C_5}$ на рис. 4.2 являются максимальными циклами графа K_5.

Рассмотрим свойство «быть кликой». Максимальной кликой будет только сам граф K_5, потому что какую бы другую клику мы не взяли, например $K_5[\{1, 2, 3, 4\}]$, её можно превратить в клику большего размера, добавив остальные вершины и рёбра: $K_5[\{1, 2, 3, 4\}] \subsetneq K_5[\{1, 2, 3, 4, 5\}] = K_5$.

Рассмотрим свойством «быть кликой чётного размера». Любая клика графа K_5 на четырёх вершинах, например $K_5[\{1, 2, 3, 4\}]$, будет максимальной кликой чётного размера, а любая клика размера два (ребро) хоть и будет обладать этим свойством, максимальной кликой чётного размера не будет.

4.4 Связность

Вершина u называется ***достижимой*** из v, если есть путь из v в u. Граф G называется ***связным***, если любая его вершина достижима из любой другой. Простая аналогия для понятия связности — города и дороги. Если V — множество городов, а E — множество дорог, то связность графа означает, что из любого города можно добраться до любого другого. Даже в реальной жизни граф автодорог может быть несвязным, в частности несвязен граф автодорог России — Калининградская область отрезана от России странами Европы.

Поэтому, при анализе дорожных сетей интересны ***компоненты связности*** — максимально связные подграфы графа G. То есть, H — компонента связности графа G, если $H \subseteq G$, H — связный граф и не существует связного подграфа $H' \subseteq G$, такого что $H \subsetneq H'$.

Очевидно, что компонента связности всегда индуцированный подграф, поэтому компоненту связности часто определяют как максимальное по включению подмножество вершин $U \subseteq V$, в котором каждая вершина достижима из любой другой. Заметим, что это определение не эквивалентно используемому нами, поскольку второе дано в терминах подграфов, а первое — в терминах подмножеств множества вершин, хотя между ними и есть взаимно однозначное соответствие: $U \mapsto G[U]$.

Любой граф является объединением его компонент связности. Изучив позже отношения эквивалентности, мы докажем, что компоненты связности либо не пересекаются, либо совпадают, то есть, если $H_1, \ldots, H_m$ — компоненты связности графа G, то

$$G = H_1 \cup H_2 \cup \ldots \cup H_m, \quad H_i \cap H_j = \varnothing \text{ при } i \neq j.$$

Пока этим фактом можно пользоваться без доказательства; попробуйте доказать это утверждение самостоятельно. Компонента связности может состоять из одной вершины; в этом случае вершина имеет степень ноль и называется ***изолированной***.

Более подробно о связности мы поговорим на следующей лекции. На этой мы докажем следующую лемму.

Лемма 1. *Пусть $G(V, E)$ — связный граф и ребро e лежит на цикле; тогда граф $G' = (V, E \setminus \{e\})$ связный. То есть, удаление ребра цикла не нарушает связность.*

Перед доказательством введём вспомогательные обозначения. Пусть P и Q — пути в графе G, x, y — вершины, лежащие на пути P, а y и z — вершины, лежащие на пути Q. Обозначим через xPy — подпуть пути P, начинающийся с вершины x и заканчивающийся в вершине y; считаем, что вершины $p_0, p_1, \ldots, p_n$ пути P упорядочены так, что $x = p_i$, $y = p_j$, $i < j$ и соответственно $xPy = p_iPp_j$ — путь на вершинах $p_i, \ldots, p_j$. Если в результате объединения путей xPy и yQz получится путь, то мы обозначаем этот путь через $xPyQz$. Это обозначение переносится и на объединение нескольких путей, а если пути P и Q имеют единственную общую вершину — общий конец, то путь, получившийся их объединением обозначим через PQ.

Доказательство леммы 1. Пусть ребро e лежит в подграфе-цикле C. Обозначим через $Q \subseteq C$ подграф-путь, получающийся из C удалением ребра e (с сохранением его концов). Зафиксируем все пути между всеми парами вершин перед удалением e и рассмотрим путь P с началом в вершине w и концом в вершине z.

Если ребро e не лежит на пути P, то после его удаления этот путь не пострадает. Если же e лежит на пути, то превратим этот путь в другой путь с помощью пути Q. Упорядочим вершины P; пусть вершина x — первая общая вершина путей P и Q (ближайшая к w, возможно сама w), а y — последняя общая вершина путей P и Q (ближайшая к z, быть может сама z). Вершины x и y определены, потому что пути P и Q имеют хотя бы две общие вершины — концы ребра e.

Докажем, что $wPxQyPz$ — путь, соединяющий вершины w и z, и не проходящий через ребро e. Действительно, пути wPx и xQy не имеют общих вершин, кроме x, поскольку иначе в пути P нашлась бы вершина ближе к w, чем x, которая была бы общая с путём Q, что противоречит выбору x; симметрично пути xQy и yPz не имеют общих вершин, кроме y (иначе нашлась бы общая вершина ближе к z, чем y); пути wPx и yPz не имеют общих вершин, поскольку это непересекающиеся подпути пути P.

Итак, мы доказали, что после удаления ребра e в графе по-прежнему останутся пути между всеми парами вершин, т. е. граф останется связным. □

Лекция 5

Графы II. Деревья и раскраски

План:

1. Связность. Теорема «#компонент связности $\geqslant |V| - |E|$».
2. Деревья. Теорема об эквивалентности четырёх свойств.
3. Расстояние между вершинами, диаметр графа. Диаметр любого связного графа не превосходит $|V| - 1$.
4. Двураскрашиваемый граф. Граф двураскрашиваемый тогда и только тогда, когда нет циклов нечётной длины.
5. Сюжет о трёх попарно знакомых или попарно незнакомых.
6. Маршруты и замкнутые маршруты. Между двумя вершинами графа есть путь, если между ними есть маршрут.

* Эйлеровы маршруты

Продолжим изучение связности.

Теорема 2. *Обозначим через C число компонент связности графа $G(V, E)$. Для него справедливо неравенство*

$$C \geqslant |V| - |E|.$$

Доказательство. Зафиксируем количество вершин в графе $|V|$ и докажем утверждение индукцией для графов с числом рёбер $|E|$ от 0 до $|V|$.

База: При $|E| = 0$ число компонент связности совпадает с числом вершин: $C = |V|$.

Шаг: Пусть для $|E| = n$ утверждение доказано. Граф с $(n + 1)$-м ребром получается из некоторого графа с n рёбрами добавлением ребра. Для графа на n

рёбрах неравенство выполняется, а добавленное ребро либо соединит две вершины из одной компоненты связности, что не уменьшит C, но уменьшит $|V| - |E|$, либо соединит вершины из разных компонент, что уменьшит и левую и правую часть неравенства на 1. В каждом из случаев, верное неравенство перейдёт в верное. □

Замечание 1. В доказательстве мы оставили читателю в качестве упражнения следующее утверждение: при добавлении в граф ребра, соединяющего вершины разных компонент связности, эти компоненты связности объединяются в одну компоненту.

Неравенство из теоремы довольно слабое: в случае, если в графе много рёбер, то число в правой части неравенства становится отрицательным, а в левой — всегда положительно. Основным применением этой теоремы служит следующее следствие.

Следствие 2. *Если граф связный, то* $|E| \geqslant |V| - 1$.

Следствие устанавливает нижнюю границу на число рёбер в связном графе, что приводит нас к определению графов специального вида — деревьев.

5.1 Деревья

Важным частным случаем неориентированных графов являются деревья. Для деревьев достаточно много эквивалентных определений, поэтому, мы будем называть граф деревом, если он удовлетворяет любому из следующих свойств:

(1) Минимально связный граф (т. е. при удалении любого ребра граф становится несвязным).

(2) Связный граф, в котором $|E| = |V| - 1$.

(3) Ациклический связный граф (связный граф без циклов).

(4) Граф, любая пара вершин которого связана единственным путём.

Докажем эквивалентность этих свойств.

Теорема 3. *Свойства (1)-(4) эквивалентны.*

Доказательство. Докажем эквивалентность, установив импликации по цепочке:

$$(2) \Rightarrow (1) \Rightarrow (3) \Rightarrow (4) \Rightarrow (2).$$

Импликаци $(2) \Rightarrow (1)$ установлена в следствие 2: из связного графа с $|V| - 1$ ребром нельзя удалить ребро без нарушения связности.

Установим импликацию $(1) \Rightarrow (3)$, воспользовавшись контрапозицией, т. е. докажем $\neg(3) \Rightarrow \neg(1)$. Отрицание условия (3) означает, что граф несвязен или имеет цикл, а условия (1), что граф или несвязен или связен, но не минимиально. Если граф несвязен, то импликация $\neg(3) \Rightarrow \neg(1)$ выполняется, поэтому сосредоточимся на случае связного графа, который содержит цикл. В лемме 1 мы установили, что

при удалении ребра из цикла в связном графе, граф остаётся связным, т. е. граф до удаления ребра был не минимально связным.

Также докажем (3) $\Rightarrow$ (4), доказав контрапозицию $\neg(4) \Rightarrow \neg(3)$. Если выполнено условие $\neg(4)$ и между какой-то парой вершин нет ни одного пути, то граф несвязынй и справедливо условие $\neg(3)$. Осталось доказать следующую лемму.

Лемма 2. *Если между вершинами w и z графа есть два различных пути P и Q, то граф содержит цикл.*

Доказательство леммы. Заметим, что $w \neq z$ (из вершины в себя ведёт единственный путь — длины 0). Если w и z единственные общие вершины путей P и Q, то склеив два пути получится цикл. Если же нет, допустим, что у путей P и Q существуют общие вершины x и y, такие что у путей xPy и xQy нет общих вершин, кроме концов, и один из путей имеет длину хотя бы 2. В этом случае, при объединении путей xPy и xQy, получится цикл.

Чтобы найти x будем двигаться вдоль путей P и Q от w к z, пока не встретится первая несовпадающая вершина. Такое обязательно случится, иначе пути совпадают. Будем считать, что несовпадающая вершина u лежит на пути P и $u \neq z$ (иначе поменяем P и Q местами).

Выберем вершину перед u в качестве x. В качестве y выберем первую после x общую вершину путей xPz и xQz. Таким образом получим, что пути xQy и xPy не имеют общих вершин кроме концов, и длина пути P хотя бы 2, и в объединении они дают цикл. □

Осталось доказать импликацию (4) $\Rightarrow$ (2). Проведём доказательство индукцией по числу вершин в графе. <u>База:</u> при $|V| = 1$ в графе нет рёбер и в вершину в себя есть единственный путь длины 0. <u>Шаг.</u> Пусть утверждение верно для всех графов на n вершинах и пусть G — произвольный граф, удовлетворяющий условию (4), в котором $V(G) = n + 1$. Выберем в G самый длинный путь P, конец которого обозначим через z.

Докажем от противного, что вершина z имеет степень 1. Допустим, что у вершины z есть ещё сосед x, кроме предшествующей ей вершины y на пути P. Если вершина x не лежит на пути P, то к пути P можно добавить ребро zx и сделать его длиннее — противоречие с выбором P. Если же x лежит на пути P, то, поскольку $x \neq y$, в графе есть два пути, соединяющие вершины x и z: xPz и ребро xz, что противоречит условию (4).

Удалив вершину z из графа G получим связный граф G' на n вершинах, для которого справедливо предложение индукции: $|E(G')| = n - 1$, поскольку между любой парой вершин G' существует единственный путь. Вернув z на место, получаем, что мы увеличили на единицу и число вершин и число рёбер графа G', а потому доказали, что $|E(G)| = |V(G)| - 1$; шаг индукции доказан. □

При доказательстве импликации (4) $\Rightarrow$ (2) мы доказали следующее свойство деревьев.

Утверждение 3. *В любом дереве, более чем с одной вершиной, есть хотя бы две вершины степени* 1.

5.2 Расстояние между вершинами. Диаметр графа

Расстоянием $\rho(u, v)$ между двумя вершинами u и v в связном графе называют длину кратчайшего пути между ними. ***Диаметром*** связного графа G называют число $\operatorname{diam}(G) = \max\limits_{u,v \in V} \rho(u, v)$, а также путь такой длины.

Утверждение 4. *Диаметр связного графа не превосходит* $|V| - 1$.

Доказательство. Действительно, если в путь входят все вершины графа, то его длина равна $|V| - 1$, а каждая вершина графа входит в любой путь не более одного раза. □

5.3 Правильные раскраски

Зафиксируем в роли цветов числа от 1 до k. ***Раскраска*** графа — это функция f, которая ставит в соответствие каждой вершине графа некоторый цвет, т. е. $f(u) \in \{1, \ldots, k\}$. Раскраска f называется ***правильной***, если концы всех рёбер покрашены в разные цвета, т. е. для каждого ребра $\{u, v\}$ справедливо $f(u) \neq f(v)$. Минимальное число цветов, в которое можно правильно раскрасить граф G называется ***хроматическим числом*** и обозначается через $\chi(G)$. Ясно, что для каждого связного графа более чем с одной вершиной $\chi(G) \geqslant 2$. Сосредоточимся на случае двураскрашиваемых графов; под ***k-раскрашиваемым*** графом мы понимаем граф, для которого существует правильная раскраска в k цветов.

Теорема 4. *Граф G является двураскрашиваемым тогда и только тогда, когда в нём нет циклов нечётной длины.*

Доказательство. Докажем сначала, что в двураскрашиваемом графе нет циклов нечётной длины. По контрапозиции, это условие равносильно тому, что если в графе есть цикл нечётной длины, то его нельзя раскрасить в два цвета. Это утверждение легко проверить. Если правильная раскраска есть, то в силу симметрии можно считать, что первая вершина цикла покрашена в цвет 1, тогда вторая вершина покрашена в цвет 2 и так далее, то есть каждая нечётная вершина будет покрашена в цвет 1, а каждая чётная — в цвет 2. Тогда последняя вершина цикла будет покрашена в тот же цвет, что и первая, что невозможно.

Докажем теперь, что если в графе нет циклов нечётной длины, то он двураскрашиваемый. Для этого построим раскраску. Выберем в каждой компоненте связности по вершине, которую назовём центром, и покрасим её в цвет 2; все вершины на расстоянии 1 от неё покрасим в цвет 1, все вершины на расстоянии 2 — в цвет 2 и так далее: вершины на чётном расстоянии от центра покрасим в цвет 2, а на нечётном в цвет 1. Предположим, что в результате этой процедуры получилась неправильная раскраска. Это означает, что у некоторого ребра $\{u, v\}$ концы были покрашены в один цвет, а это произошло, если расстояния от центра c некоторой компоненты до вершин u и v имеют одинаковую чётность. Пусть P — кратчайший путь от центра до u, а Q — кратчайший путь от центра до v и w самая дальняя от центра их общая вершина (быть может сам центр, если других общих вершин нет). Заметим, что w не совпадает ни с u, ни с v: иначе мы получили бы,

что расстояния до u и v отличаются на единицу; по этой же причине ребро $\{u, v\}$ не лежит ни на одном из этих путей. Пути cPw и cQw имеют одинаковую длину; в противном случае один из этих путей можно было бы заменить на более короткий другой и сократить длину пути до u или v.

Отсюда мы получаем, что пути wPu и wQv пересекаются только по вершине w и их длины имеют одинаковые чётности (от длин одинаковой чётности отнимается расстояние от c до w). Объединив эти пути и добавив к ним ребро $\{u, v\}$ получим цикл нечётной длины, что приводит нас к противоречию. □

5.4 Сюжет про трёх попарно знакомых или попарно незнакомых

Рассмотрим ещё один занимательный пример, подобный примеру о рукопожатиях. Если выбрать любые 6 человек в аудитории, то среди них окажется или три попарно знакомых человека или три попарно незнакомых. Переведём это утверждение на язык теории графов.

Утверждение 5. *В графе на шести вершинах есть или клика размера* 3 *или независимое множество размера* 3.

Доказательство. Возьмём любую вершину. Тогда она либо смежна с какими-то тремя из оставшихся, либо не смежна с какими-то тремя из оставшихся. Эти случаи симметричны, давайте считать, что выполняется первый случай и вершина a смежна с вершинами b, c и d. Если среди b, c и d есть хотя бы одно ребро, то концы этого ребра вместе с a являются вершинами клики размера 3; если же рёбер между ними нет, то $\{b, c, d\}$ — независимое множество размера 3. □

5.5 Эйлеровы маршруты

Маршрутом в графе G называется последовательность вершин $v_0, v_2, \ldots, v_n$, такая что $n \geqslant 0$ и $\{v_i, v_{i+1}\} \in E(G)$ для $0 \leqslant i \leqslant n-1$. Обратим внимание, что второе условие применимо только в случае, когда в маршруте больше одной вершины, а потому любая последовательность из единственной вершины считается маршрутом. ***Длина маршрута*** — это число рёбер, соединяющих вершины маршрута; оно совпадает с n. Маршрут называется ***замкнутым***, если $v_0 = v_n$. Будем говорить. что ребро $\{x, y\} \in E(G)$ ***лежит*** на маршруте, если для некоторого i выполняется $\{v_i, v_{i+1}\} = \{x, y\}$.

Очевидно, что если в графе есть путь из вершины x в вершину y, то в нём есть и маршрут между этими вершинами. Обратное тоже справедливо.

Утверждение 6. *Если в графе есть маршрут между вершинами x и y, то есть и соединяющий их путь.*

Доказательство. Пусть $v_0, \ldots v_n$, $v_0 = x$, $v_n = y$ — самый короткий маршрут из x в y. Докажем, что в нём нет повторяющихся вершин. Действительно, если $v_i = v_j, i < j$, то удалив из маршрута все вершины с v_{i+1} до v_j получим маршрут

короче. Значит в выбранном маршруте нет повторяющихся вершин, соседние вершины маршрута по определению соединены ребром, отсюда получаем, что граф на вершинах маршрута со всеми рёбрами маршрута является графом-путём. □

Теория графов зародилась в городе Кёнигсберг (ныне Калининград) со следующей задачи. Математик Леонард Эйлер любил гулять и поставил цель: обойти все острова Кёнигсберга, пройдя по каждому мосту ровно один раз, вернувшись в конце прогулки в начало пути. План местности из работы Эйлера приведён на рис. 5.1.

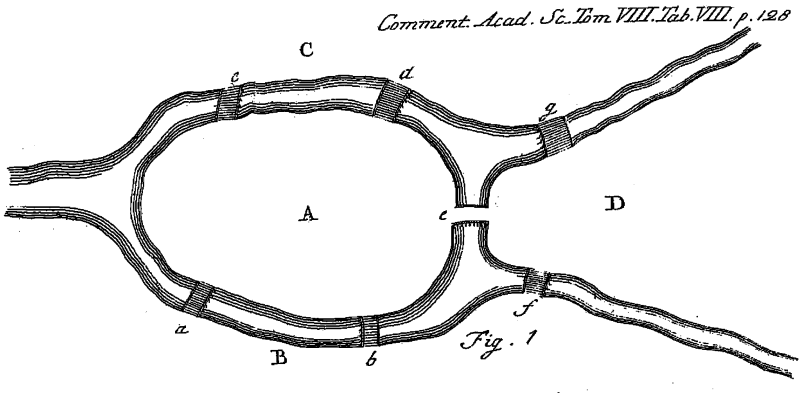


Рис. 5.1. Мосты Кёнигсберга

Переведём задачу на язык графов. Эйлера интересовало, существует ли на графе ниже замкнутый маршрут, который содержит все рёбра графа ровно один раз. Маршрут, который содержит все рёбра графа ровно один раз назовём ***эйлеровым маршрутом***. Обратим внимание, что граф из оригинальной задачи содержит кратные рёбра, которые мы не рассматриваем.

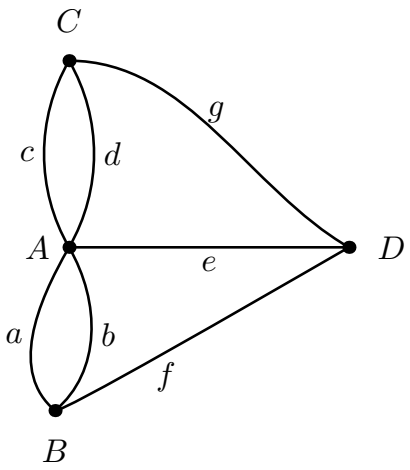


Рис. 5.2. Граф для задачи о мостах Кёнигсберга.

Эйлер обнаружил, что нужного маршрута не существует и доказал простой критерий существования такого маршрута.

Теорема 5. *Связный граф G содержит замкнутый эйлеров маршрут тогда и только тогда, когда степень каждой вершины чётна.*

Доказательство. Докажем сначала, что если в графе есть вершина нечётной степени, то в нём нет замкнутого Эйлерова маршрута. Пусть x — вершина нечётной степени. Если в графе есть замкнутый эйлеров маршрут, то циклически сдвинув его вершины легко добиться, чтобы маршрут начинался и заканчивался в x. Заметим, что каждый раз, когда вершина x встречается в середине маршрута, в маршруте встречаются сразу два ребра x: в x нужно сначала зайти по одному ребру, а потом выйти по другому. Поскольку x является первой и последней вершиной маршрута, то на его концах также задействуется два ребра: первое на первом выходе из x, а второе — при последнем возврате. Значит, что в во всём маршруте участвовало только какое-то чётное число рёбер, смежных x, а всего их нечётное число — получается, что хотя бы одно ребро, смежное с x, в маршрут не попало.

Перейдём теперь к доказательству основной импликации: из чётности всех степеней связного графа следует существование замкнутого эйлерова маршрута. Пусть $R = r_0, r_1, \ldots, r_m$ — маршрут максимальной длины, в который каждое ребро графа G входит не более одного раза. В случае, если таких маршрутов несколько, то выберем любой из них. Установим два свойства такого маршрута, которые и приведут к доказательству существования искомого маршрута.

Свойство 1. $r_0 = r_m$. Предположим противное. Повторяя те же аргументы, что и в первом абзаце, получим, что между r_1 и r_m встречается чётное число рёбер, смежных с r_m, и ещё одно ребро, $r_{m-1}r_m$ встречается в конце маршрута. Итого на маршруте R лежит нечётное число рёбер, смежных с r_m, а поскольку степень вершины r_m чётна, то есть ещё хотя бы одно ребро $r_m x$, которое не лежит на маршруте R. Добавив вершину x в конец маршрута получим маршрут большей длины, что противоречит выбору R.

Свойство 2. Маршрут R содержит все рёбра графа G. Допустим противное. Пусть некоторое ребро xy графа G не лежит на R. Рассмотрим путь из r_0 в x, который существует в силу связности G и найдём на нём первое ребро, которое не лежит на R, если оно есть. Обозначим его через $r_i z$. Если такого ребра нет, то вершина x лежит на R, а потому обозначим $r_i = x$ и $z = y$. По свойству 1, маршрут R замкнутый, а потому сдвинем его циклически так, чтобы он начинался и заканчивался в вершине r_i. Добавив в конец получившегося маршрута вершину z получим маршрут, длиннее R, который также содержит каждое ребро не более одного раза, что приводит нас к противоречию выбора R. □

Лекция 6

Двудольные графы, паросочетания и функции

План:

1. Двудольные графы и паросочетание
2. Теорема Холла
3*. Доказательство теоремы Холла [7]
4. Функции (область определения, множество значений, образ, полный прообраз)
5. Отображения (всюду определённые функции): инъекции, сюръекции, биекции
6. Отображения и задача о назначениях

Граф $G(V, E)$ называется ***двудольным***, если существует разбиение множества V на подмножества L и R ($V = L \cup R$, $L \cap R = \varnothing$), такие что у каждого ребра один конец лежит в L, а другой в R, т. е. между вершинами из L нет рёбер, как и между вершинами из R. Множества L и R называют ***долями графа***. Вообще говоря, в двудольном графе может быть несколько разбиений множества вершин на доли. Когда мы говорим о двудольном графе $G(L \cup R, E)$ мы подразумеваем, что разбиение на доли зафиксировано.

Двудольный граф $G(L \cup R, E)$ называется ***полным***, если $E = \{\{l, r\} \mid l \in L, r \in R\}$, то есть полный двудольный граф содержит всевозможные (для двудольного графа) рёбра. Полный двудольный граф с долями из m и n вершин обозначают через $K_{m,n}$. Граф звезда — это граф $K_{1,n}$.

Лемма 3. *Граф двудольный тогда и только тогда, когда он двураскрашиваемый.*

Двудольные графы часто встречаются в задачах о назначениях. Пусть L — множество процессоров, а R — множество задач; между вершинами l и r есть ребро, если процессор l может решить задачу r (в многопроцессорных системах это зависит от многих параметров). Возникает естественная задача: распределить задачи по процессорам наиболее эффективно, то есть добиться того, чтобы как можно больше задач было решено в единицу времени.

Эта задача известна как задача об оптимальном паросочетании. ***Паросочетание*** — это множество рёбер M, в котором ни одна пара рёбер не имеет общего конца. Ясно, что паросочетание с наибольшим числом рёбер определяет оптимальное распределение задач по процессорам. В идеальном случае, получиться распределить все задачи по всем процессорам, такое паросочетание называется ***совершенным***; формально $M \subseteq E$ — совершенное паросочетание, если каждая вершина графа смежна с некоторым ребром из M; вершины, смежные с рёбрами из M, называют ***покрытыми*** M.

Пример 9. На рисунке 6.1 приведено совершенное паросочетание в двудольном графе с долями $\{a, b, c, d\}$ и $\{w, x, y, z\}$. Рёбра, входящие в паросочетание выделены жирным.

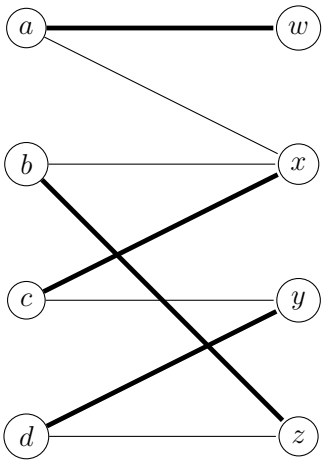


Рис. 6.1. Совершенное паросочетание

Замечание 2. Данные определения о паросочетаниях справедливы для произвольных графов, хотя мы и сосредоточимся на двудольных.

Есть простое условие, которое позволяет проверить, есть ли в двудольном графе совершенное паросочетание. Чтобы его сформулировать, нам потребуется понятие окрестности. Обозначим ***множество соседей*** (***окрестность***) вершины v через

$$N(v) = \{u \mid \{u, v\} \in E\}.$$

Обобщим понятие соседей на подмножество вершин: соседи множества S — это вершины, смежные хотя бы с одной вершиной из S и не лежащие в S:

$$N(S) = \left(\bigcup_{v \in S} N(v) \right) \setminus S.$$

Заметим, что в случае двудольного графа справедливо $N(S) = \bigcup\limits_{v \in S} N(v)$ при $S \subseteq L$ или $S \subseteq R$.

Теорема 6 (теорема Холла о свадьбах)**.** *В двудольном графе с долями L и R существует совершенное паросочетание тогда и только тогда, когда $|L| = |R|$ и для любого подмножества $S \subseteq L$ справедливо*

$$|N(S)| \geqslant |S|.$$

Доказательство теоремы Холла

Для доказательства теоремы нам потребуются следующие вспомогательные определения. Пусть в двудольном графе с долями L и R зафиксировано паросочетание M. Путь положительной длины называется ***чередующимся***, если он начинается в непокрытой M вершине из доли L и рёбра пути чередуются между не принадлежащими M и принадлежащими M. На рис. 6.2.а любой путь, начинающийся из вершины 1 является чередующимся.

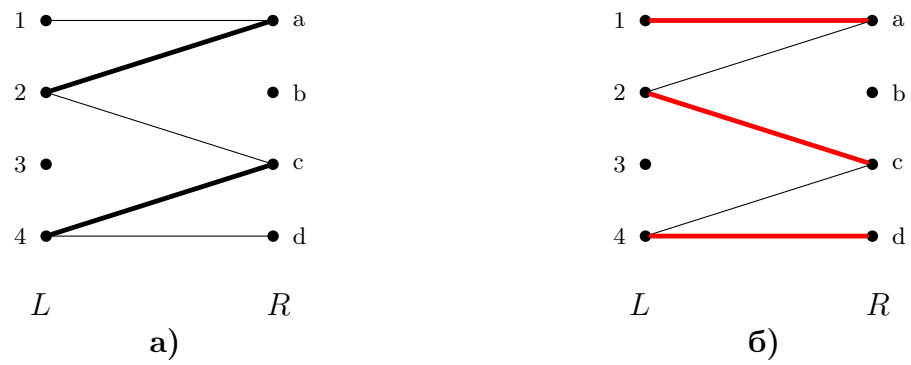


Рис. 6.2. Чередующие и увеличивающие пути

Из определения чередующегося пути вытекает, что каждое ребро пути из доли L в долю R не принадлежит M, а каждое ребро из доли R в долю L принадлежит (мы считаем, что рёбра пути направлены от первой вершины к последней). Заметим также, что чередующийся путь чётной длины заканчивается в доле L, а нечётной длины — в доле R. Также ясно, что в путях нечётной длины непокрытых рёбер (не входящих в M) на единицу больше, чем покрытых. Эти наблюдения приводят нас к понятию увеличивающего пути.

Чередующийся путь называется ***увеличивающим***, если он заканчивается в непокрытой вершине. Из сказанного выше ясно, что это возможно только в случае, если последняя вершина лежит в правой доле, и таким образом путь будет нечётной длины. Смысл названия состоит в том, что при замене в паросочетании M покрытых рёбер увеличивающего пути на непокрытые, получается паросочетание M' большего размера. Перед доказательством этого утверждения обратимся сначала к интуиции. Как видно из рис. 6.2.б при замене рёбер размер паросочетания M увеличивается на единицу и при этом в нём остаются покрыты все ранее покрытые вершины

увеличивающего пути, и помимо них добавляются ещё первая и последняя вершины пути. Докажем, что это будет выполняться для произвольного увеличивающего пути.

Утверждение 7. *Замена в паросочетании M всех покрытых рёбер увеличивающего пути P на непокрытые приводит к паросочетанию M' большего размера.*

Доказательство. Все вершины пути P, покрытые M останутся покрытми, потому что они являются концами как чётных, так и нечётных рёбер P. К покрытым вершинам добавятся ещё первая и последняя вершины P, ранее не покрытые в M (по определению увеличивающего пути). Никакие рёбра M, не входящие в P, при этом не меняются, а количество покрытых M вершин увеличивается, значит увеличивается и количество рёбер в новом паросочетании M'. □

Замечание 3. Определения чередующегося и увеличивающегося путей справедливы и для путей длины 1. В этом случае путём является непокрытое паросочетанием ребро (во втором случае оба конца обязательно не покрыты). Проверьте, что рассуждения выше и ниже проходят и для этого вырожденного случая.

Перейдём к доказательству теоремы Холла. Необходимость условия очевидна: если совершенное паросочетание существует, то у каждой вершины в левой доли есть уникальный сосед справа (поставленный в соответствие паросочетанием); таким образом, соотношение $|N(S)| \geqslant |S|$ справедливо для любого подмножества $S \subseteq L$.

Сосредоточимся на доказательстве достаточности условия теоремы. Мы сведём его к доказательству леммы.

Лемма 4. *Если M — не совершенное паросочетание в двудольном графе, для которого выполняется условие теоремы Холла, то существует паросочетание M', такое что $|M'| > |M|$.*

Объясним, почему лемма эквивалентна достаточности условия. Предположим, что в графе не существует совершенного паросочетания и возьмём паросочетание M максимального размера. Тогда из леммы получим, что существует паросочетание M' большего размера, что противоречит выбору M.

Доказательство леммы 4. Поскольку M не совершенное паросочетание, в левой доле существует непокрытая им вершина a. Обозначим через A множество всех вершин L, достижимых из a чередующимися путями, а через $B \subseteq R$ — множество всех предпоследних вершин этих путей.

Покажем, что размеры множеств A и B совпадают. Каждая вершина из A лежит на каком-то чередующемся пути из вершины a и в силу чётности, последнее ребро этого пути принадлежит паросочетанию M. Поскольку для каждой вершины $a' \in A$ существует уникальная вершина $b' \in R$, такая что $\{a', b'\} \in M$, то b' и есть предпоследняя вершина чередующегося пути, а вершин b' ровно столько же, сколько и a'.

Зафиксируем множество $S = \{a\} \cup A$ и рассмотрим множество $N(S)$. Ясно, что $B \subseteq N(S)$ и $|N(S)| \geqslant |A| + 1 > |B|$, таким образом существует вершина $b \in N(S) \setminus B$.

Поскольку $b \in N(S)$, то b лежит на чередующемся пути из a: либо b смежна с a, либо b смежна с некоторой вершиной $v \in A$; поскольку v — конец некоторого чередующегося пути из a, добавив к этому пути ребро $\{v, b\}$ получим чередующийся путь; ребро $\{v, b\}$ не покрыто M, поскольку покрыто ребро ведущее из B в v.

Вершина b не покрыта паросочетанием M: иначе из b вело бы ребро в множество A и вершина b принадлежала бы множеству B по его определению.

Таким образом, b лежит на чередующемся пути из a и не покрыта M, а значит путь из a в b — увеличивающий. Отсюда по утверждению 7 получаем, что в графе есть паросочетание M' большего размера.

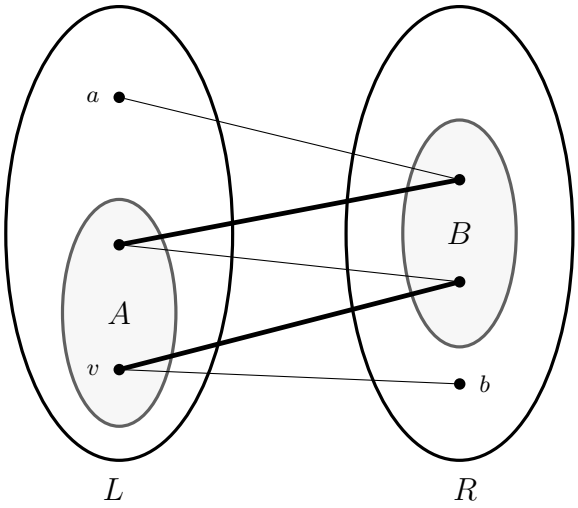


Рис. 6.3. Иллюстрация к доказательству леммы 4

□

6.1 Функции

С помощью двудольных графов легко ввести основные понятия, связанные с функциями. Начнём с картинок. На рис. 6.4 проиллюстрирована функция, а на рис. 6.5 — не функция.

Чтобы разобраться, что отличает функцию от нефункции приведём определение функции. Неформально, ***функция*** — это закон, который ставит в соответствие элементам множества X элементы множества Y; каждому элементу $x \in X$ поставлен в соответствие не более, чем один элемент y из множества Y. Используемые нами функции называются ***частично определёнными*** или ***частичными***, поскольку не обязательно для каждого $x \in X$ определён соответсвующий ему $y \in Y$.

Говорить о функции более правильно с помощью ориентированных графов. В ориентированном графе рёбра — это упорядоченные пары вершин, т. е. на рёбрах есть стрелки. По аналогии с неориентированным графом введём понятия степени вершин и (множества) соседей для ориентированного графа. Поскольку рёбра

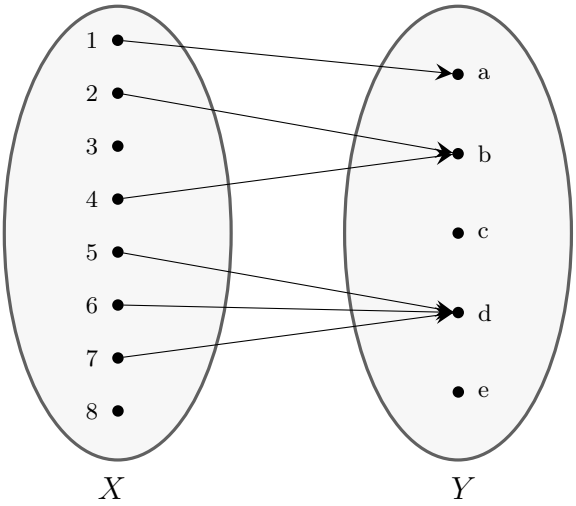


Рис. 6.4. Функция f

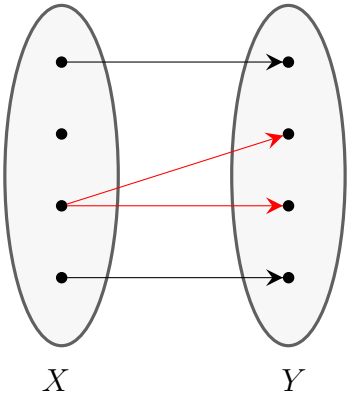


Рис. 6.5. Не функция

имеют направление, то мы разделяем исходящую степень $d_+(v)$ (число вершин достижимых из v по одному ребру) и входящую степень $d_-(v)$ (числу вершин, из которых за один шаг по ребру можно добраться до v), а также множества правых и левых соседей $N_+(v)$ и $N_-(v)$; определения соседей переносятся на подмножества вершин также, как и в случае неориентированного графа.

Из определения функции ясно, что она представима с помощью ориентированного двудольного графа (в левой доли все левые концы рёбер), в котором исходящая степень каждой вершины левой доли не больше единицы; граф при этом может быть и бесконечным.

Обозначим через f множество рёбер графа, задающего функцию f из X в Y; тогда $(x, y) \in f$ означает, что $f(x) = y$. Пусть $G(X \cup Y, f)$ — граф, для функции f.

Введём с помощью двудольного графа базовые определения, связанные с функциями и разберёмся с ними на примере функции на рис. 6.4. Рекомендуем читателю самостоятельно сделать упражнение ниже, прежде чем переходить к ответу.

- ***Областью определения*** $\mathrm{Dom}(f) \subseteq X$ называется подмножество вершин с исходящей степенью 1 (подмножество X, на котором определена функция f).
- ***Множеством значений*** $\mathrm{Range}(f) \subseteq Y$ называется подмножество вершин с входящей степенью больше 0 (подмножество Y всевозможных значений f).
- ***Образом*** $f(A)$ множества $A \subseteq X$ называют множество значений, которые принимает f на подмножестве A; на языке графов — это множество правых соседей $N_+(A)$:

$$f(A) = \{y \mid \exists x \in A : f(x) = y\} = N_+(A).$$

- ***Полным прообразом*** $f^{-1}(B)$ множества $B \subseteq Y$ называют множество элементов X, значение функции на которых лежит в B; на языке графов — это

множество левых соседей $N_-(B)$:

$$f^{-1}(B) = \{x \mid \exists y \in B : f(x) = y\} = N_-(B).$$

Если $f(x) = y$, то элемент y называют ***образом*** элемента x, а элемент x называют ***прообразом*** элемента y. Из определений следует, что $\mathrm{Dom}(f) = f^{-1}(Y)$, а $\mathrm{Range}(f) = f(X)$.

Функцию на рис. 6.4 также можно задать с помощью обозначения $x \mapsto y$:

$$f : 1 \mapsto a, \quad 2 \mapsto b, \quad 4 \mapsto b, \quad 5 \mapsto d, \quad 6 \mapsto d, \quad 7 \mapsto d.$$

Упражнение 4. Найдите для функции f на рис. 6.4 область определения, множество значений, образ множества $A = \{1, 2, 3, 4\}$ и полный прообраз множества $B = \{b, c, d\}$.

Ответ: $\mathrm{Dom}(f) = \{1, 2, 4, 5, 6, 7\}$, $\mathrm{Range}(f) = \{a, b, d\}$, $f(A) = \{a, b\}$, $f^{-1}(B) = \{2, 4, 5, 6, 7\}$.

Замечание 4. Не следует путать обозначение f^{-1} с обозначением для обратной функции. Полный прообраз определён для произвольной функции. Одинаковые обозначения неслучайны: если полный прообраз каждого элемента y содержит не более одного элемента x, то функция f обратима, и обратная к ней функция g определяется как $g(y) = x$, где $x \in f^{-1}(y)$.

6.2 Отображения

В случае $\mathrm{Dom}(f) = X$, функция f называется ***всюду определённой*** или ***отображением***[1]. Запись $f : X \to Y$ означает, что f всюду определена.

Среди отображений выделяют следующие три важных вида.

Определение 1. Отображение $f : X \to Y$ называется ***инъекцией***, если $f(x) \neq f(x')$ при $x \neq x'$. В терминах графа, это означает, что входящая степень каждого $y \in Y$ не превосходит единицу.

Примеры инъекции и не инъекции приведены на рис. 6.6.

Определение 2. Отображение $f : X \to Y$ называется ***сюръекцией***, если у каждого элемента y существует прообраз, т. е. $\mathrm{Range}(f) = Y$ или что то же самое $\forall y \in Y\ \exists x \in X : f(x) = y$. В терминах графа, это означает, что входящая степень каждого $y \in Y$ больше нуля.

Примеры сюръекции и не сюръекции приведены на рис. 6.7.

Определение 3. Отображение $f : X \to Y$ называется ***биекцией***, если оно является инъекцией и сюръекцией.

Из определения биекции ясно, что каждая вершина X и Y имеет единичную исходящую и входящую степень соответственно. То есть биекция устанавливает взаимно однозначное соответствие между множествами X и Y. Пример биекции приведён на рис. 6.8.

[1] В разных областях математики под термины функции и отображения резервируют разные понятия; будьте внимательны при чтении математической литературы.

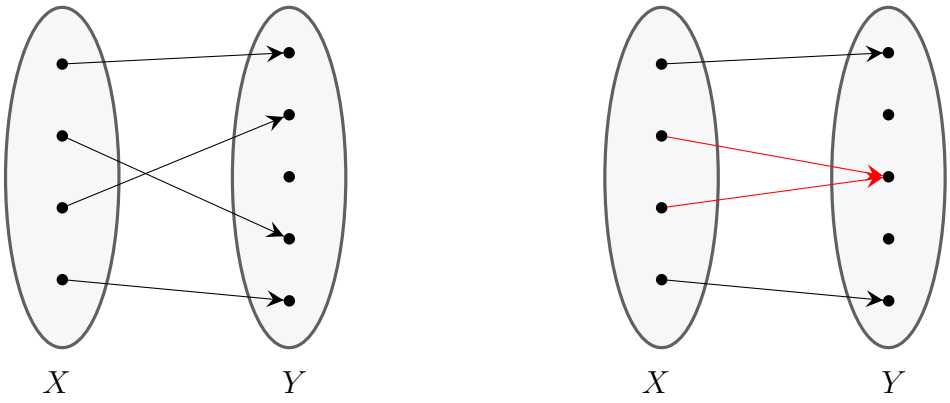


Рис. 6.6. Инъекция и не инъекция

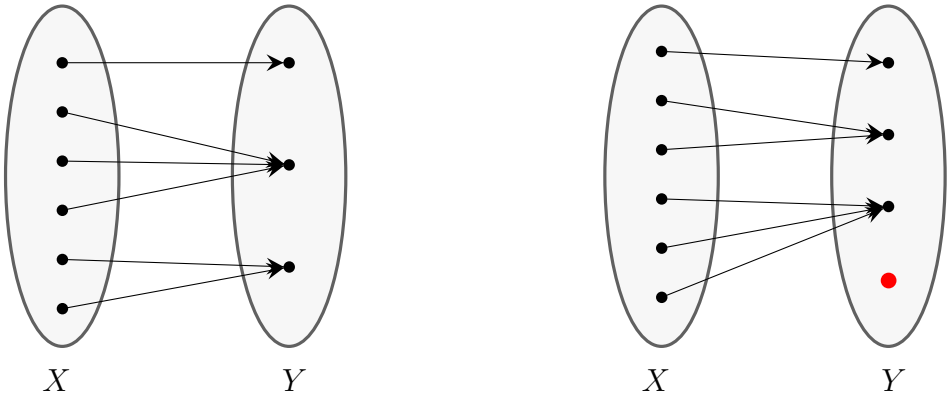


Рис. 6.7. Сюръекция и не сюръекция

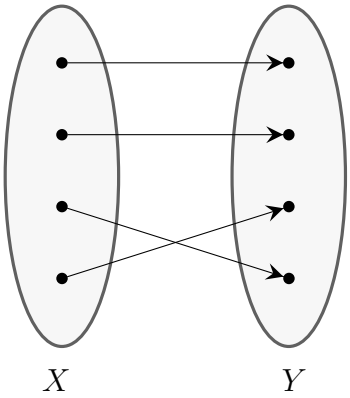


Рис. 6.8. Биекция

6.3 Функции и задача о назначениях

Вернёмся к задаче о распределении задач по процессорам. В ней задан двудольны граф, который задаёт ограничения на назначения задач процессорам. Любой подграф этого графа, задающий функцию, ставит в соответствии каждому процессору ровно одну задачу, но разным процессорам могут быть назначены одинаковые задачи.

Если каждому процессору назначено по задаче, то эта функция — отображение. Если вместе с этим разным процессорам назначены разные задачи, то эта функция — инъекция. Если каждому процессору назначено по задаче и каждая задача назначена какому-то процессору, то эта функция — сюръекция. В идеальном случае, каждому процессору назначена своя задача и распределены все задачи: в этом случае функция — биекция.

В случае, если не каждому процессу получается назначить задачу, то вместо того, чтобы считать, что задана не всюду определённая функция f из L в R можно считать, что задано отображение $f : A \to R$, где $A = \mathrm{Dom}(f) \subseteq L$. Это удобно, чтобы пользоваться понятиями инъекции, сюръекции и биекции.

Итак, задача о назначении (она же о поиске максимального паросочетания) сводится к поиску инъекции с самой большой областью определения (её размер равен как числу процессоров, так и числу распределённых задач). В идеале все задачи распределены, в этом случае найдена сюръекция; если же $|L| = |R|$ и все задачи распределены по всем процессорам, то найдено лучшее решение — биекция, которая соответствует совершенному паросочетанию.

Лекция 7

Комбинаторика I. Правила суммы и произведения

План:

1. Отображения и подсчёты
2. Правило суммы
3. Правило произведения. Биекция с декартовым произведением множеств
4. Число слов длины n над алфавитом из k символов
5. Перестановки
6. Подсчёт количества слов длины k с разными буквами. Размещения
7. Число сочетаний. Количество k-элементных подмножеств n-элементного множества
8. Подсчёты с кратностью: сколько различных слов можно составить из слова «Математика»?

9*. Дискретная вероятность

Комбинаторика — раздел математики, изучающий дискретные объекты, такие как (конечные) множества, функции, графы, и другие объекты, с которыми нам ещё предстоит познакомиться. Перечислительная комбинаторика сосредоточена на задачах о перечислении и подсчёте. Предмет перечислительной комбинаторики может на первый взгляд показаться нелепым: если у нас есть конечное множество объектов и нам нужно подсчитать их количество, то давайте просто выпишем все объекты и занумеруем их!

Объясним в чём возникает проблема. Допустим вы хотите найти число двоичных слов длины 10. Их 1024 и выписывать в ряд все двоичные слова будет довольно утомительно, а если перейти от десяти к тысяче, то уже слабо реалистично. С перечислением тоже не всё бывает просто. Допустим, нужно перечислить все 5-элементные подмножества 100-элементного множества. Это реально сделать: их всего 75287520, но как их перебрать? Перечислять все подмножества стоэлементного множества плохая идея — их число огромно:

$$126765060022822940149670320537 6.$$

Для подсчёта и перебора комбинаторика использует связь между различными дискретными объектами. Например, как подсчитать число всех подмножеств n-элементного множества? Вместо того, чтобы считать явно, докажем, что их столько же, сколько и двоичных слов длины n; из двоичной системы счисления мы знаем, что последних 2^n.

Утверждение 8. *Число подмножеств n-элементного множества совпадает с числом двоичных слов длины n.*

Доказательство. Занумеруем элементы множества от 1 до n и поставим каждому подмножеству S в соответствие двоичное слово w_S по следующему правилу: если $i \in S$, то i-й бит w_S равен 1, иначе 0.

Мы построили отображение из подмножеств n-элементного множества в слова. Это отображение инъекция: если $S_1 \neq S_2$, то существует элемент $i \in S_1 \triangle S_2$, такой что в одном из слов w_{S_1}, w_{S_2} на i-м месте будет стоять 1, а в другом 0. Ясно, что это отображение сюръекция — по каждому двоичному слову легко восстановить подмножество. Мы доказали, что установили биекцию между подмножествами n-элементного множества и двоичными словами длины n, а значит их количество совпадает. □

7.1 Отображения и подсчёты

При доказательстве утверждения 8 мы немного забежали вперёд, апеллировав к очевидности того, что если между конечными множествами A и B есть биекция, то число их элементов совпадает. В этом разделе мы заполним лакуны.

Допустим у нас есть два конечных множества A и B и нам интересно узнать, в каком из них больше элементов (или установить, что элементов поровну). Вовсе не обязательно перечислять все элементы. Например, если множество A — множество студентов, а B — множества стульев, то, чтобы узнать мощность какого множества больше, достаточно предложить студентам занять стулья. Если остались свободные стулья, то больше стульев, если остались стоящие студенты, то больше студентов. В первом случае была построена инъекция из множества студентов в множество стульев, а во втором — инъекция из множества стульев в множество студентов. Если студентов и стульев оказалось поровну, то была установлена биекция. Формализуем этот пример.

Лемма 5. *Пусть A и B — конечные непустые множества.*

- $|A| \leqslant |B| \iff$ *существует инъекция из A в B;*
- $|A| \geqslant |B| \iff$ *существует сюръекция из A в B;*
- $|A| = |B| \iff$ *существует биекция из A в B.*

Пример 10. Чего больше: разбиений числа 12 на 4 слагаемых или его разбиений на слагаемые не превосходящие 4? Разбиения, отличающиеся перестановкой слагаемых будем считать одинаковыми, такие как $5+3+2+2$ и $2+3+2+5$.

Формализуем задачу. ***Разбиением*** положительного целого числа n на слагаемые называется последовательность чисел $x_1 \geqslant x_2 \geqslant \ldots \geqslant x_m$, такая что $x_i \in \mathbb{N}_1$ и $n = x_1 + x_2 + \ldots + x_m$; поскольку порядок слагаемых неважен, мы фиксируем в формальном определении представителя для каждого разбиения (со слагаемыми упорядоченными по невозрастанию). Для ответа на вопрос поставим каждому разбиению в соответствие картинку, которая называется диаграммой Юнга. На рис. 7.1 приведены диаграммы Юнга для разбиений $5+3+2+2$ и $4+4+2+1+1$:

Рис. 7.1. Диаграммы Юнга

Число клеток в i-ой строке диаграммы Юнга совпадает с i-ым слагаемым. Таким образом мы установили биекцию между разбиениями и диаграммами Юнга. Число клеток в диаграмме совпадает с разбиваемым числом n; количество строг совпадает с числом слагаемых, а количество столбцов равно максимальному слагаемому.

Определим на диаграммах Юнга операцию транспонирования, по аналогии с транспонированием матриц: геометрически транспонирование состоит из последовательных применений поворота налево на 90° и отражения относительно оси Ox.

Диаграммы на рис. 7.1 получаются из друг друга транспонированием. Очевидно, что транспонирование задаёт биекцию между множеством диаграмм Юнга с n клетками с не более, чем k стоками, и множеством диаграмм Юнга с n клетками и не более чем k столбцами. Установив биекцию между этими множествами мы получили, что в них одинаковое число элементов, что означает, что количество разбиений числа n на не более чем k слагаемых совпадает с количеством разбиений числа n на слагаемые, не превосходящие k.

7.2 Правило суммы

Правило суммы становится очевидным после изучения теории множеств. Оно гласит, что если конечные множества A и B не пересекаются, то мощность их объединения совпадает с суммой мощностей:

$$|A \cup B| = |A| + |B|, \text{ если } A \cap B = \varnothing.$$

В общем случае, из диаграмм Эйлера-Венна легко получить, что

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Пример 11. Найдём с помощью правила суммы количество четырёхзначных чисел. Обозначим через A множество четырёхзначных чисел, а через B множество чисел от 1 до 999. Тогда объединение $A \cup B$ состоит из чисел от 1 до 9999. Ясно, что $|B| = 999$ и $|A \cup B| = 9999$, отсюда $|A| = |A \cup B| - |B| = 9000$.

7.3 Правило произведения

Правило произведения формулируется на естественном языке следующим образом. Если есть n объектов первого типа и после выбора любого объекта первого типа можно выбрать m объектов второго типа, то всего есть $n \times m$ способов последовательно выбрать первый и второй объект.

Пример 12. Найдём количество двузначных чисел с помощью правила произведения. Старший разряд числа может быть любой цифрой от 1 до 9 а младший — цифрой от 0 до 9. Таким образом есть 9 способов выбрать старший разряд и после каждого выбора есть 10 способов выбрать младший разряд. Итого двухзначных чисел 90 по правилу произведения.

Правило произведения легко обобщается по индукции на k последовательных выборов. Если объект первого типа можно выбрать n_1 способами, после чего второй объект можно выбрать n_2 способами и т. д. (k-й объект можно выбрать n_k способами), то выбрать последовательно k объектов можно $n_1 \times n_2 \times \ldots \times n_k$ способами.

Так аналогично подсчёту двухзначных чисел можно подсчитать количество трёхзначных чисел и k-значных чисел. Оставляем общий случай читателю в качестве упражнения.

При первом знакомстве с правилом произведения создаётся впечатление, что последовательные выборы объектов должны быть независимы, но это не так. Рассмотрим следующий пример.

Пример 13. Для дежурства на перемене учителю нужно выбрать из класса, в котором 20 человек, двух дежурных — старшего дежурного и его помощника. Требуется найти число способов это сделать.

Из применения правила произведения получаем ответ: $20 \times 19 = 380$. Отметим, что после выбора старшего дежурного всегда будет 19 вариантов выбора его помощника, но множества этих вариантов отличаются друг от друга — в них всегда отсутствует только выбранный старший дежурный.

Процесс последовательно выбора можно проиллюстрировать с помощью дерева

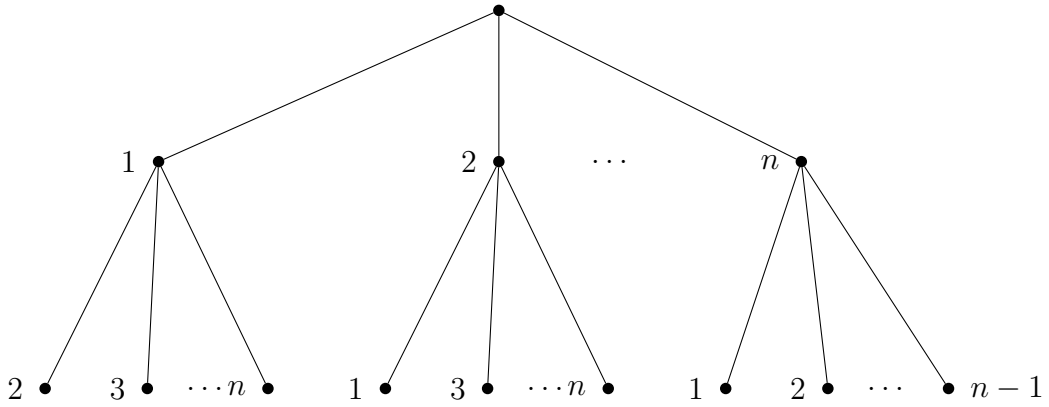


Рис. 7.2. Дерево последовательных выборов

Какая операция на множествах соответствует правилу произведения? Эта операция называется декартовым произведением. ***Декартовым произведением*** множеств X и Y называется множество упорядоченных пар элементов из множеств X и Y соответственно:

$$X \times Y = \{(x, y) \mid x \in X,\ y \in Y\}.$$

Имя Декарта знакомо читателю со школьной скамьи по декартовой системе координат. Декартова система координат иллюстрирует геометрически декартово произведение множеств $\mathbb{R} \times \mathbb{R}$ (рис. 7.3).

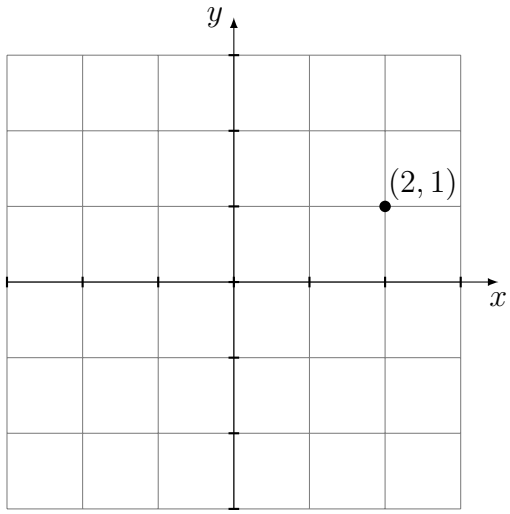


Рис. 7.3. Декартово произведение $\mathbb{R} \times \mathbb{R}$

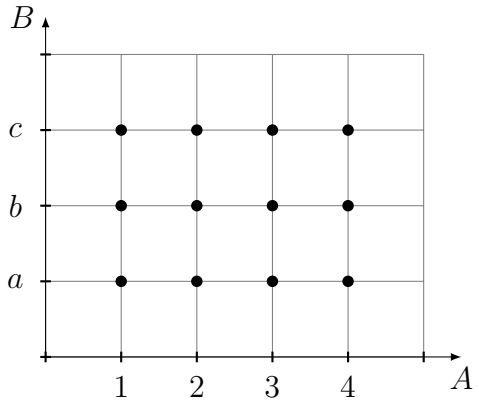


Рис. 7.4. Декартово произведение конечных множеств

Проиллюстрируем графически декартово произведение множеств $A = \{1, 2, 3, 4\}$ и $B = \{a, b, c\}$. Из рис. 7.4 видно, что мощность множества $A \times B$ есть произведение мощностей множеств A и B: $|A \times B| = |A| \times |B|$. На рис. 7.5 декартово произведение проиллюстрировано в виде дерева, которое можно рассматривать как дерево последовательных выборов.

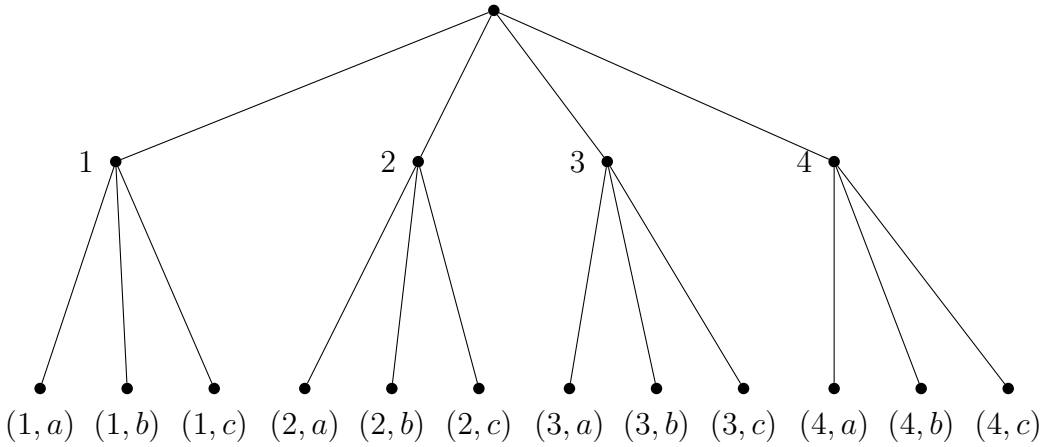


Рис. 7.5. Иллюстрация декартова произведения через дерево

Формально правило произведения можно определить через биекцию с декартовым произведением: пусть $A = \{1, \ldots, n\}$, $B = \{1, \ldots, m\}$; если существует биекция между множествами C и $A \times B$, то $|C| = n \times m$.

Подсчёт слов, перестановок, размещений и подмножеств

Слово — это конечная последовательность символов, которые в свою очередь определяются как элементы конечного множества — ***алфавита***. В зависимости от задачи, под алфавитом из k символов часто удобно понимать множество $[k]_0 = \{0, 1, \ldots, k-1\}$ или $[k]_1 = \{1, \ldots, k\}$. Первое полезно при работе с системами счисления. Выше мы установили, что слов длины n над двоичным алфавитом 2^n. В общем случае, число слов длины n над k-ичным алфавитом равно k^n по правилу произведения: на первую позицию можно поставить любую из k букв, после чего любую из k букв можно поставить на второе место и так вплоть до n-ой.

Чтобы лучше понять природу последующих объектов, будем решать естественные задачи, в которых они потребуются.

Пример 14. В мешке есть n разных шаров и нужно расставить их все на полку в ряд. Сколько есть способов это сделать?

Воспользуемся правилом произведения: на первое место на полке можно поставить любой из n шаров, после него можно поставить любой $n-1$ из оставшихся, затем $n-2$, и так вплоть до последнего, для которого остаётся единственный вариант.

Итак, получилось произведение чисел от n до 1, которое называют ***факториалом*** числа n и обозначают

$$n! = n \times (n-1) \times (n-2) \times \ldots \times 1,$$

при этом считается, что $0! = 1$.

Решение задачи можно свести к подсчёту слов специального вида. Слова над алфавитом $[n]_1$ длины n, в которых все символы разные. Такие слова называются ***перестановками***. Разобравшись с примером, мы установили, что число перестановок есть $n!$.

Пример 15. В мешке есть n разных шаров и нужно расставить k из них на полку в ряд. Сколько есть способов это сделать?

Решение этой задачи почти не отличается от предыдущей, только нужно в произведении остановиться после k-ого шага.

В терминах слов, требуется найти количество слов длины k над алфавитом $[n]_1$, в которых все символы разные. Такие слова называются ***размещениями*** и для их количества используют обозначения

$$A_n^k = n \times (n-1) \times (n-2) \times \ldots \times (n-k+1) = \frac{n!}{(n-k)!}.$$

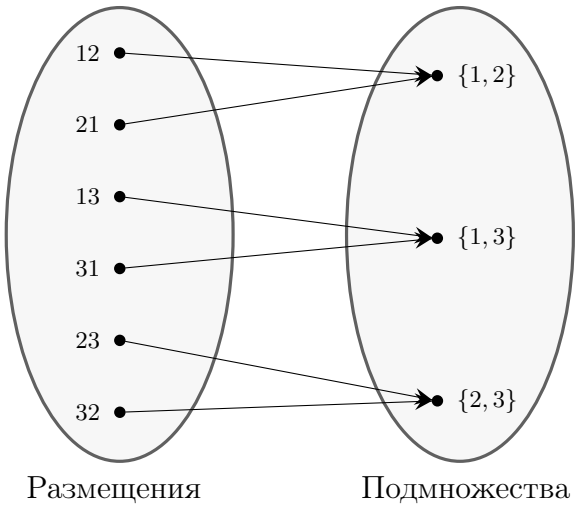
Обозначение A_n^k называется числом размещений («A» от arrangement).

Пример 16. В мешке есть n разных шаров и нужно положить k из них в другой мешок. Сколько есть способов это сделать?

Начнём с формализации задачи. Требуется найти количество k-элементных подмножеств n-элементного множества. Ясно, что любой мешок можно получить так: расставить k шаров на полке, а потом сложить их в мешок. Этим рассуждением мы определили отображение из множества размещений (слов над $[n]_1$ длины k с различными буквами) в множество подмножеств (k-элементнных подмножеств n-элементного множества):

$$f : w_1 w_2 \ldots w_k \mapsto \{w_1, w_2, \ldots, w_k\}$$

Построим двудольный граф для этого отображения для случая $n = 3$, $k = 2$:



Из примера видно, что в каждое подмножество идёт по две стрелочки. Это верно и в общем случае: при любых $n, k > 0$ и размер полного прообраза любого подмножества одинаковый и равен $k!$. Для доказательства этого факта воспользуемся первым примером: всего есть $k!$ вариантов расставить k шаров из мешка на полке.

Теперь у нас есть всё необходимое, чтобы решить задачу. Нам известно число размещений $\frac{n!}{(n-k)!}$, помимо этого нам известно, что построенное отображение f — сюръекция (в любое подмножество ведёт хотя бы одна стрелка) и мы установили, что размер полного прообраза любого элемента равен $k!$. Таким образом мы получили, что если $\binom{n}{k}$ — число k-элементных подмножеств n-элементного множества, то $\binom{n}{k} \times k! = \frac{n!}{(n-k)!}$ отсюда получаем, что

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}.$$

Число $\binom{n}{k}$ называют ***числом сочетаний***.

МАТЕМАТИКА

Перейдём к следующей задаче. Сколько разных не обязательно осмысленных слов можно получить переставляя буквы слова МАТЕМАТИКА?

Предположим, на минуту, что все буквы разные:

$$\mathrm{M_1A_1T_1EM_2A_2T_2ИКА_3},$$

тогда всего слов будет 10!. Давайте теперь отождествим буквы М, определив отображения из множество слов, в которых все буквы разные, в множество слов, в которых М одинаковые стиранием индексов. У каждого слова с одинаковыми М будет ровно два прообраза, поэтому таких слов будет $\frac{10!}{2}$. Отождествим теперь буквы А стиранием индексов: определим отображения из слов, в которых одинаковые только М, в слова, в слова, в которых одинаковые и М и А (стиранием индексов). У каждого слова с одинаковыми М и А одинаковое число прообразов: 3! — столько различных способов расставить индексы у трёх А:

$$\mathrm{MA_1T_1EMA_2T_2ИКА_3}, \quad \mathrm{MA_1T_1EMA_3T_2ИКА_2}, \quad \mathrm{MA_2T_1EMA_1T_2ИКА_3}\ldots$$

Продолжая те же рассуждения, отождествим буквы Т и получим что искомое количество есть

$$\frac{10!}{2!3!2!}.$$

Решим теперь задачу другим способом. Будем строить слово, последовательно расставляя буквы. Всего есть 10 позиций. Выберем сначала 2 позиции под букву М:

$$\underset{1}{\mathrm{M}}\ \underline{\ }\ \underline{\ }\ \underline{\ }\ \underline{\ }\ \underset{5}{\mathrm{M}}\ \underline{\ }\ \underline{\ }\ \underline{\ }\ \underline{\ }\ \underline{\ }.$$

Сколько способов это сделать? Занумеруем все позиции; чтобы расставить буквы М нужно выбрать ровно две из них, причём порядок позиций не важен, то есть число расстановок совпадает с числом двухэлементных подмножеств десятиэлементного множества — $\binom{10}{2}$. Расставим теперь буквы А: для них осталось 8 свободных позиций и нужно расставить 3 буквы А, аналогично рассуждениям с М это можно сделать $\binom{8}{3}$ способами. Таким образом, число способов расставить М и А по правилу произведения равняется $\binom{10}{2} \times \binom{8}{3}$. Продолжая рассуждения получаем, что число способов сделать это есть

$$\binom{10}{2} \times \binom{8}{3} \times \binom{5}{2} \times \binom{3}{1} \times \binom{2}{1} \times \binom{1}{1}.$$

Обратим внимание читателя, что последние три множителя есть просто 3!, потому что $\binom{n}{1} = n$, то есть после расстановки повторяющихся букв, число способов выбрать позицию под одну букву совпадает с числом оставшихся позиций и правило произведения даёт, что число способов расставить 3 неповторяющиеся буквы есть ни что иное как 3!.

Решив комбинаторную задачу двумя способами мы установили справедливость следующей формулы:

$$\binom{10}{2} \times \binom{8}{3} \times \binom{5}{2} \times \binom{3}{1} \times \binom{2}{1} \times \binom{1}{1} = \frac{10!}{2!3!2!1!1!1!};$$

мы добавили в знаменатель правой части для симметрии три $1!$ — по $1!$ на каждую неповторяющуюся букву.

Решив задачу про слово МАТЕМАТИКА мы доказали следующее утверждение.

Утверждение 9. *Пусть $k_1 + k_2 + \ldots + k_m = n$. Число слов длины n над алфавитом из m символов, в которых первая буква алфавита встречается k_1 раз, вторая — k_2 раз, i-я — k_i раз есть*

$$\frac{n!}{k_1!k_2!\ldots k_m!} = \binom{n}{k_1} \times \binom{n-k_1}{k_2} \times \binom{n-k_1-k_2}{k_3} \times \ldots \times \binom{n-k_1-\ldots-k_{m-1}}{k_m}.$$

Это утверждение пригодится нам дальше.

Лекция 8

Комбинаторика II. Биномиальные коэффициенты

Лекция была прочитана близко к изложению во второй главе [1].

План:

1. Количество путей по узлам клеток (вправо и вверх) из (0,0) в (i,j) есть $\binom{i+j}{i}$.
2. Треугольник Паскаля и его свойства
   - симметрия
   - возрастание биномиальных коэффициентов к середине
   - оценка $\binom{2n}{n} > \dfrac{2^{2n}}{2n+1}$.
3. Бином Ньютона и биномиальные коэффициенты
   - рекуррентное соотношение
   - сумма биномиальных коэффициентов и её комбинаторный смысл
   - знакопеременная сумма биномиальных коэффициентов
4. Комбинаторные доказательства
   - рекуррентное соотношение на биномиальные коэффициенты в треугольнике Паскаля;
   - задача о командире и солдатах: $n \times 2^{n-1} = \sum\limits_{k=1}^{n} k\binom{n}{k}$;
   - $\sum\limits_{k=1}^{n} \binom{n}{k}^2 = \binom{2n}{n}$;
5. Метод точек и перегородок

- Число решений уравнения $x_1 + x_2 + \ldots + x_k = n$ в неотрицательных целых числах есть $\binom{n+k-1}{k-1}$ (Формула Муавра).

6. Числа Фибоначчи

7*. Числа Каталана (доказательство явной формулы)

Лекция 9

Комбинаторика III. Формула включений-исключений

План:

1. Доказательство формулы включений-исключений (через характеристические функции)
2. Примеры
   - Количество чисел от 1 до 100 не делящихся ни на 3, ни на 5, ни на 7
   - Связь со знакопеременной суммой биномиальных коэффициентов
   - Задача о счастливых билетах
3. Подсчёт числа отображений (всюду определённых функций), частичных функций, инъекций, биекций
4. Подсчёт сюръекций
5. Комбинаторные объекты. Смысл обозначений 2^A для множества всех подмножеств и Y^X для множества отображений из X в Y.
6. Принцип Дирихле

9.1 Правило суммы и формула включений-исключений

Напомним, что правило суммы гласит, что если множества A и B не пересекаются (т. е. $A \cap B = \varnothing$), то $|A \cup B| = |A| + |B|$. Или ещё его формулируют так: «Если

есть n способов выбрать первый объект и m способов выбрать второй объект, и эти способы независимы, то всего есть $n+m$ способов выбрать первый или второй объект.»

Что же делать, если множества A и B пересекаются? Ясно, что тогда

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Пример 17. Сколько чисел от 1 до 100 не делятся ни на 2, ни на 3?

Давайте посчитаем сколько чисел делятся на 2 или на 3; обозначив через $D_k \subseteq \{1, \ldots, 100\}$ подмножество из чисел, делящихся на k получаем:

$$|D_2 \cup D_3| = |D_2| + |D_3| - |D_2 \cap D_3| = |D_2| + |D_3| - |D_6|.$$

То есть $50 + 33 - 16 = 67$. Отсюда, ни на 2 ни на 3 не делятся 33 числа.

Нарисовав диаграммы Эйлера-Венна для трёх множеств нетрудно получить, что

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

Эти формулы помогают посчитать число элементов во всех множествах, когда известно количество элементов во всевозможных пересечениях этих множеств. В общем случае формула-включения исключений устроена так (в формулах $[n] = \{1, 2, \ldots, n\}$):

$$|A_1 \cup A_2 \cup \ldots \cup A_n| = \sum_{i=1}^{n} |A_i| - \sum_{1 \leqslant i < j \leqslant n} |A_i \cap A_j| + \ldots + (-1)^{m+1} \sum_{\substack{S \subseteq [n], \\ |S|=m}} \left| \bigcap_{i \in S} A_i \right| + \ldots$$

или более компактно

$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{\substack{S \subseteq [n], \\ S \neq \varnothing}} (-1)^{|S|+1} \left| \bigcap_{i \in S} A_i \right|. \tag{1}$$

Сумма в последней формуле берётся по всем непустым подмножествам $[n]$; так, если $S = \{1, 2, 4\}$, то соответствующее слагаемое имеет вид $(-1)^4 \times |A_1 \cap A_2 \cap A_4|$.

Формулу включений-исключений можно прямолинейно доказать по индукции. Однако дело это неблагодарное, поэтому мы приведём другое доказательство — через характеристические функции.

Характеристические функции

Зафиксируем юнивёрсум U. Функция $\chi_A(x)$ называется характеристической функцией множества $A \subseteq U$, если

$$\chi_A(x) = \begin{cases} 1, & \text{если } x \in A; \\ 0, & \text{если } x \notin A. \end{cases}$$

Характеристические функции на первый взгляд ничем не отличаются от предикатов. Разница лишь в том, что предикаты принято относить к логическим

формулам, а арифметические действия с ними проводить не принято (если сложить две истинны, то это не считается более убедительной истиной). С помощью характеристической функции легко выразить мощность множества:

$$|A| = \sum_{x \in U} \chi_A(x).$$

Для доказательства формулы включений-исключений нам потребуется выразить характеристическую функцию пересечения, объединения и дополнения множеств через элементарные множества:

$$\chi_{A \cap B}(x) = \chi_A(x) \times \chi_B(x);$$

$$\chi_{A \cup B}(x) = \chi_A(x) + \chi_B(x) - \chi_A(x) \times \chi_B(x);$$

$$\chi_{\bar{A}}(x) = 1 - \chi_A(x).$$

Упражнение 5. Проверьте корректность этих формул.

Доказательство формулы включений-исключений

Основным ингредиентом доказательства является обобщённый закон де Моргана:

$$\bigcup_{i=1}^{n} A_i = \overline{\bigcap_{i=1}^{n} \bar{A}_i},$$

который мы переведём на язык характеристических функций (для удобства записи опустим аргументы у функций):

$$\chi_{\bigcup\limits_{i=1}^{n} A_i} = 1 - \chi_{\bigcap\limits_{i=1}^{n} \overline{A_i}} = 1 - (1 - \chi_{A_1}) \times (1 - \chi_{A_2}) \times \ldots \times (1 - \chi_{A_n}).$$

Раскроем скобки в выражении $(1 - \chi_{A_1})(1 - \chi_{A_2}) \ldots (1 - \chi_{A_n})$ и проанализируем получившееся выражение аналогично анализу для бинома Ньютона. Ясно, что среди слагаемых есть 1; чтобы её получить нужно взять в качестве сомножителей 1 из каждой скобки; также для каждого i в выражение войдёт слагаемое $-\chi_{A_i}$: чтобы получить его нужно взять $-\chi_{A_i}$ из i-ой скобки, а из каждой оставшейся скобки взять единицу. Продолжая рассуждения получаем, что чтобы получить слагаемое $\prod\limits_{i \in S} \chi_{A_i} = \chi_{A_{i_1}} \times \chi_{A_{i_2}} \times \ldots \times \chi_{A_{i_{|S|}}}$ нужно взять из каждой скобки с номером i из множества S слагаемое $-\chi_{A_i}$, а из остальных скобок взять 1; при этом коэффициент перед получившимся произведением будет $(-1)^{|S|}$. Итак, мы получили формулу

$$\chi_{\bigcup\limits_{i=1}^{n} A_i} = 1 - \left(1 + \sum_{\substack{S \subseteq [n], \\ S \neq \varnothing}} (-1)^{|S|} \prod_{i \in S} \chi_{A_i}\right) = \sum_{\substack{S \subseteq [n], \\ S \neq \varnothing}} (-1)^{|S|+1} \chi_{\bigcap\limits_{i \in S} A_i},$$

в последнем переходе мы перешли от произведения характеристических функций к характеристической функции пересечения множеств.

Просуммировав обе части по всем $x \in U$ получим требуемую формулу (1). Обратим внимание, что в результате суммирования левая часть формулы будет иметь вид

$$\sum_{x \in U} \chi_{\bigcup\limits_{i=1}^{n} A_i}(x) = \left| \bigcup_{i=1}^{n} A_i \right|,$$

а в правой части в результате суммирования отдельного слагаемого получится следующее

$$\sum_{x \in U} (-1)^{|S|+1} \chi_{\bigcap\limits_{i \in S} A_i}(x) = (-1)^{|S|+1} \times \left| \bigcap_{i \in S} A_i \right|.$$

9.2 Задача о счастливых билетах

В качестве примера применения формулы включений-исключений приведём задачу о счастливых билетах. В XX веке в СССР и России билеты на автобусы имели номера, состоящие из шести цифр. Билеты нужно было компостировать (дырявить специальным устройством в автобусе), а если сумма первых трёх цифр совпадала с суммой последних трёх, то билет считался счастливым, и его полагалось на удачу съесть, что приводило к казусам, если в автобус заходил контролёр. Отсюда возникла естественная для комбинаторики задача — подсчитать количество счастливых билетов.

Обозначим первые три цифры как a_1, a_2, a_3, а последние через b_1, b_2, b_3. Сделаем замену $a_4 = 9 - b_1, a_5 = 9 - b_2, a_6 = 9 - b_3$. Получаем, что билет счастливый тогда и только тогда, когда

$$a_1 + a_2 + a_3 + a_4 + a_5 + a_6 = 27, \quad 0 \leqslant a_i \leqslant 9. \tag{2}$$

Сходу возможно неясно, как подсчитать все решения, удовлетворяющие (2). Давайте для начала найдём все решения, в которых $a_i \geqslant 0$. Это классическая задача Муавра, всего таких решений $\binom{32}{5}$. Среди этих решений есть и те, которые не годятся, давайте их вычтем.

Ясно, что не годятся те решения, в которых хотя бы одна переменная принимает значение 10 или больше. Выберем одну (плохую) переменную, которой присвоим значение 10, после чего распределим оставшиеся 17 единиц по всем переменным (включая плохую); число способов сделать последнее совпадает с числом решений уравнения $\sum\limits_{i=1}^{6} a_i = 17$, т. е. $\binom{22}{5}$ и всего есть 6 способов выбрать плохую переменную; по правилу произведения получаем, что плохих вариантов $6 \times \binom{22}{5}$.

Но ответ $\binom{32}{5} - 6 \times \binom{22}{5}$ неверный: при вычитании мы два раза учли случай когда первая переменная была выбрана плохой, а вторая получила 10 или больше единиц при последующем распределении и наоборот. Поэтому к этой разности нужно добавить произведение $\binom{6}{2} \times \binom{12}{5}$ — число способов выбрать две переменные, которым сразу выдадут 10 единиц, после чего распределить оставшиеся 7 единиц по 6 переменным. Поскольку трёх плохих переменных быть уже не может, получаем итоговый ответ:

$$\binom{32}{5} - 6 \times \binom{22}{5} + \binom{6}{2} \times \binom{12}{5}.$$

Давайте посмотрим на эти рассуждения внимательно. Вроде бы мы использовали формулу включений-исключений, но это произошло неявно. Какие же множества нужно было выбрать, чтобы ею воспользоваться?

С помощью формулы включений-исключений мы посчитали плохие случаи. Обозначим через A_i множество решений уравнения 2, в котором для i-ой переменной условие не выполняется; формально

$$A_i = \{(a_1, a_2, \ldots, a_6) \mid a_i \geqslant 10,\ a_1 + a_2 + \ldots + a_6 = 27,\ a_j \geqslant 0,\ \text{при } 1 \leqslant j \leqslant 6\}.$$

Чтобы подсчитать все плохие решения достаточно вычислить мощность объединения множеств $|A_1 \cup A_2 \cup \ldots \cup A_6|$, для которого справедливо

$$|A_1 \cup A_2 \cup \ldots \cup A_6| = |A_1| + |A_2| + \ldots + |A_6| - |A_1 \cap A_2| - |A_1 \cap A_3| - \ldots - |A_5 \cap A_6| =$$

$$= 6 \times |A_1| - \binom{6}{2} \times |A_1 \cap A_2|.$$

Тройные и последующие пересечения множеств в формулу не входят, так как пересечение любых трёх множеств (среди A_i) пусто. В силу симметрии в каждом множестве A_i одинаковое число элементов, равно как и в попарных пересечениях. Убедитесь, что в решении задачи о счастливых билетах выше мы действительно использовали эту формулу.

9.3 Подсчёт функций

Как и раньше, обозначим через $[m]$ — m-элементное множество; для определённости $[m] = \{1, 2, \ldots, m\}$. Найдём число отображений $f : [m] \to [n]$ (из m-элементного множества в n-элементное множество). Напомним, что отображения — это всюду определённые функции, этот факт обозначается с помощью стрелочки между множествами. Для значения $f(1)$ годится любое из n чисел, равно как для $f(2)$ и так вплоть до $f(m)$; по правилу произведения получаем, что число отображений есть n^m. Каждое отображение можно закодировать как слово длины m над алфавитом из n символов. Эта кодировка пригодится нам дальше.

Найдём теперь число частичных функций из $[m]$ в $[n]$. В отличие от отображений, частичные функции не обязательно всюду определены, поэтому возможно, что у каких-то элементов множества $[m]$ нет образов. Эту задачу легко свести к предыдущей, добавив к множеству $[n]$ элемент $n+1$ и построить для каждой функции f из $[m]$ в $[n]$ эквивалентное отображение $g : [m] \to [n+1]$, которое принимает значение $n+1$ во всех точках, в которых функция f не определена, а в остальных точках принимает то же значение, что и f. Итак, мы установили биекцию между множеством частичных функций из $[m]$ в $[n]$ и множеством отображений из $[m]$ в $[n+1]$, таким образом число частичных функций равно $(n+1)^m$.

Перейдём теперь к подсчёту инъекций из $[m]$ в $[n]$. Вспомним, что инъекция — отображение, которое ставит в соответствие разным элементам разные значения. Мы уже обсуждали тот факт, что если существует инъекция из конечного множества A в конечное множество B, то $|A| \leqslant |B|$, поэтому при $m > n$ инъекций нет вовсе. В случае $0 < m \leqslant n$ инъекцию, как и любое отображение, можно закодировать в виде слова над n-ичным алфавитом длины m, а условие инъективности

означает, что в слове все буквы разные. Таким образом, число инъекций совпадает с числом размещений и равно $\frac{n!}{(n-m)!}$.

В случае конечных множеств, биекция является частным случаем инъекции, когда $[m] = [n]$. Таким образом число биекций равно $n!$ (при $m = n$) и совпадает с числом перестановок. Совпадение чисел размещений и перестановок с числом инъекций и биекций неслучайно. Инъекции кодируют размещения, а биекции — перестановки.

Подсчёт сюръекций

Перейдём теперь к подсчёту сюръекций. В отличие от подсчитанных выше функций, подсчитать число сюръекций уже непросто. Аналогично случаю с инъекциями необходимо, чтобы $m \geqslant n > 0$. Чтобы подсчитать сюръекции подсчитаем сначала отображения, не являющиеся сюръекциями и вычтем их число из количества всех отображений.

Напомним, что функция является сюръекцией, если прообраз каждого из элементов $[n]$ не пуст. Таким образом, функция не является сюръекцией, если в $[n]$ есть хотя бы один элемент y для которого нет подходящего $x : \forall x \in [m] : f(x) \neq y$. Обозначим через A_i — множество отображений, для которых прообраз элемента i не определён, формально

$$A_i = \{f \mid f : [m] \to [n],\ f^{-1}(i) = \varnothing\}.$$

Ясно, что все несюръекции лежат в множестве $\bigcup\limits_{i=1}^{n} A_i$.

Подсчитаем мощность $\left|\bigcup\limits_{i=1}^{n} A_i\right|$ с помощью формулы включений-исключений. Для каждого A_i число $|A_i|$ совпадает с числом отображений из m-элементного множества в $(n-1)$-элементное множество (i-й) элемент задействовать нельзя, а все остальные можно. Их число есть $(n-1)^m$. В пересечении множеств $A_i \cap A_j$ $(i \neq j)$ содержится столько же элементов, сколько отображений $[m] \to [n-2]$ (теперь нельзя задействовать ровно два элементов), а число элементов в пересечении любых k множеств из семейства $A_1, A_2, \ldots, A_n$ совпадает с числом отображений $[m] \to [n-k]$. Число способов выбрать k множеств из семейства с n-множествами есть число сочетаний $\binom{n}{k}$, отсюда получаем по формуле включений исключений

$$\begin{aligned}\left|\bigcup_{i=1}^{n} A_i\right| &= |A_1| + |A_2| + \ldots + |A_n| - |A_1 \cap A_2| - \ldots |A_{n-1} \cap A_n| + \ldots = \\ &= n \times (n-1)^m - \binom{n}{2} \times (n-2)^m + \binom{n}{3} \times (n-3)^m - \ldots = \\ &= \sum_{k=1}^{n} (-1)^{k+1} \times \binom{n}{k} \times (n-k)^m.\end{aligned}$$

Таким образом число сюръекций есть

$$n^m - \left(n \times (n-1)^m - \binom{n}{2} \times (n-2)^m + \binom{n}{3} \times (n-3)^m - \dots \right) =$$
$$= \sum_{k=0}^{n} (-1)^k \times \binom{n}{k} \times (n-k)^m.$$

9.4 О комбинаторных объектах

При подсчётах отображений специального вида мы получили, что разные комбинаторные объекты выражаются с помощью друг друга: размещения реализуются с помощью инъекций, перестановки — с помощью биекций. Сюръекции выражают размещения m разных объектов по n разным ящикам, в которых не учитывается порядок содержимого ящиков; содержимое k-го ящика — полный прообраз $f^{-1}(k)$.

Численные совпадения при подсчёте комбинаторных объектов часто влекут за собой и содержательную интерпретацию, которую часто получается раскрыть введя правильную кодировку. Вспомним, что в начале изучения комбинаторики мы кодировали любое подмножество S m-элементного множества с помощью двоичной строки длины m. Такую же кодировку имеют и отображения $[m] \to \{0,1\}$; как мы убедились в начале этой лекции каждому множеству S соответствует характеристическая функция χ_S, поэтому этот способ кодирования часто называют ***характеристическим вектором***. Таким образом, для любого подмножества S множества X существует эквивалентное отображение $\chi_S : X \to \{0,1\}$, то есть подмножества и отображения в двухэлементные множества по сути одинаковые комбинаторные объекты.

Подсчёты отображений привели к следующим устоявшимся обозначениям. Множество отображений из X в Y обозначают Y^X: в случае конечных непустых множеств, как мы убедились, их число есть $|Y|^{|X|}$. Множество всех подмножеств множества X обозначают как 2^X, поскольку ему взаимно однозначно соответствует множество отображений $[2]^X$.

Напомним также, что множество k-элементных подмножеств X обозначается $\binom{X}{k}$, что приводит к формуле

$$\bigcup_{k=0}^{|X|} \binom{X}{k} = 2^X,$$

справедливой в случае конечного множества X; для этой формулы есть эквивалентная формула для мощностей: $\sum\limits_{k=0}^{n} \binom{n}{k} = 2^n$.

9.5 Принцип Дирихле

Очевидно, что если в семье три ребёнка, то два из них одного пола. Подобное рассуждение продолжается так: если есть m голубей и их нужно рассадить по n клеткам, то в случае, если $n < m$, в некоторой клетке окажется хотя бы два

голубя. В примере с семьёй голуби — это дети, а клетки — это полы[1]. Этот принцип получил название принцип голубятни (Pigeonhole principle), а в русскоязычном математическом сообществе он известен как принцип Дирихле.

После изучения отображений и их свойств принцип Дирихле становится очевидным: он гласит всего лишь, что в случае $n < m$ не существует инъекции из m-элементного множества в n-элементное множество. Отсюда вытекает, что в случае отображения $f : [m] \to [n]$, существует элемент $y \in [n]$, в полном прообразе которого лежит хотя бы два элемента: $|f^{-1}(y)| \geqslant 2$.

[1]Мы придерживаемся здесь традиционных взглядов, и считаем что полов всего два (мужской и женский). Хотя бы у детей.

Лекция 10

Бинарные отношения и их графы. Отношения эквивалентности

План:

1. Задание бинарного отношения таблицей, двудольным графом, перечислением пар. Формальное определение бинарных отношений ($R \subseteq A \times B$).
2. Некоторые свойства
   - функциональность
   - тотальность
   - инъективность
   - рефлексивность
   - транзитивность
   - симметричность
3. Отношения эквивалентности. Примеры:
   - Рациональные числа
   - Равные и подобные треугольники
   - Равенство булевых функций
   - Неопределённые интегралы
4. **Т.:** Классы эквивалентности не пересекаются или совпадают.
5. Следствие: отношения эквивалентности взаимно однозначно соответствуют разбиениям множества на подмножества.
6. Пример использования в комбинаторике: подсчёт числа k-элементных подмножеств сводится к подсчёту числа классов эквивалентности на наборах: $(x_1, x_2, \ldots, x_k) \sim (y_1, y_2, \ldots, y_k) \iff \{x_1, x_2, \ldots, x_k\} = \{y_1, y_2, \ldots, y_k\}$.

7. Теоретико-множественные операции с отношениями. Операция обращения. Описание с помощью булевых матриц.

8. Композиция отношений. Связь с базами данных

10.1 Описания и определение бинарных отношений

Начнём с примера, который иллюстрирует что такое бинарное отношение. Представьте, что состоялась контрольная, которую писало три человека, и её результаты приведены в таблице ниже:

Ученик	1	2	3	4	5
Маша	+	−	+	−	−
Алина	+	−	−	−	+
Джон	−	−	+	−	+

Рис. 10.1. задание бинарного отношения таблицей

Считается, что каждый ученик, либо решил задачу, либо нет. Если школьник A решил задачу x, то они находятся в отношении «решил задачу». Формально, построенная нами таблица — это 0-1-матрица

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Результаты контрольной можно было бы также описать с помощью двудольного графа (рис. 10.2).

Пожалуй, проще всего было задать отношение «решил задачу» просто перечислив все пары из школьников и задач:

(Маша, 1), (Маша, 3), (Алина, 1), (Алина, 5), (Джжон, 3), (Джон, 5).

Этот способ и лежит в основе формального определения бинарного отношения. Формально ***бинарное отношение*** R между множествами A и B — это некоторое подмножество их декартова произведения:

$$R \subseteq A \times B.$$

Если $(a, b) \in R$, говорят, что элемент a ***находится в отношении*** R с элементом b.

Выше мы привели различные способы описания бинарного отношения: перечислением пар, таблицей (матрицей), двудольным графом. Переход от одного способа описания к другому бывает полезен на практике; вернёмся к примеру из лекции о

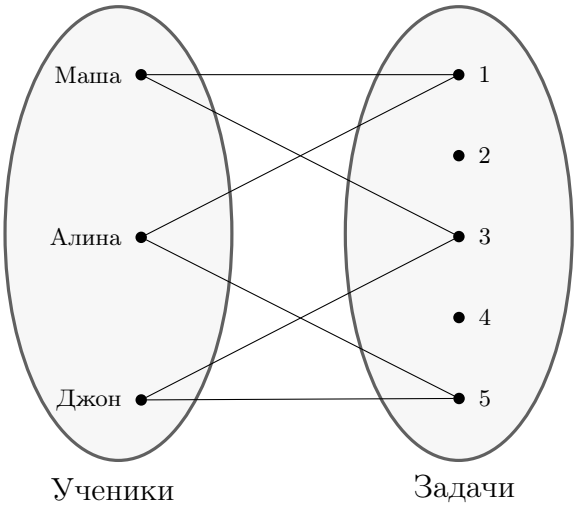


Рис. 10.2. Задание бинарного отношения двудольным графом

паросочетаниях. Представьте, что A — множество процессоров, а B — множество задач. Множество пар (a, b) задаёт отношение R — выполнимости задачи b на процессоре a. Задача состоит в выборе подмножества $f \subseteq R$, такого что в единицу времени решается максимальное число задач, при этом каждый процессор решает не более одной задачи. На языке двудольных графов это означает, что требуется найти максимальное паросочетание. Если при этом получилось задействовать все процессоры, то на языке функций это означает, что найденное отношение f — инъекция, а если при этом задействованы и все задачи, то f — биекция.

10.2 Примеры и свойства

В последнем примере предыдущего раздела мы упомянули, что функция — частный случай бинарного отношения. Если быть точнее, то формальное определение функции и даётся через бинарное отношение.

Определение 4. Бинарное отношение $f \subseteq X \times Y$ называется ***функциональным*** или ***функцией***, если из $(x, y) \in f$ и $(x, y') \in f$ следует, что $y = y'$.

В случае функций факт $(x, y) \in f$ обозначают через $f(x) = y$.

Пусть $R \subseteq X \times Y$ В случае, если $\forall x \in X \, \exists y \in Y : (x, y) \in R$, отношение R называется ***(левым-)тотальным***. Тотальное и функциональное отношение $f \subseteq X \times Y$ является всюду определённой (тотальной) функцией или отображением.

Отношение $R \subseteq X \times Y$ является ***инъективным***, если из $(x, y) \in R$ и $(x', y) \in R$ следует, что $x = x'$.

Отношение $R \subseteq X \times Y$ является ***сюръективным***, если $\forall y \in Y \, \exists x \in X : (x, y) \in R$.

Тотальные функциональные инъективные отношения задают инъекции, тотальные функциональные сюръективные — сюръекции, тотальные функциональные инъективные и сюръективные (всё одновременно) — биекции.

Бинарные отношения знакомы вам со школьной скамьи. Символы $<, >, \leqslant, \geqslant, =, \neq$ задают бинарные отношения. Чтобы не было путаницы, когда мы говорим об отношениях, заданных математическими символами, мы будем окружать их скобками: например, формально $(\leqslant) \subseteq \mathbb{R} \times \mathbb{R}$ можно задать через множество как

$$(\leqslant) = \{(x, y) \mid x \leqslant y\}.$$

Читатель может быть недовольным, что мы задали отношение $(\leqslant)$, используя его же внутри; мы сделали это лишь для наглядности. Формально это множество можно было бы описать и с помощью декартовой плоскости: в него входят все точки лежащие на прямой $y = x$ и все точки выше этой прямой.

Мы будем использовать устоявшееся ***обозначение*** xRy, равносильное $(x, y) \in R$; вы к нему уже хорошо привыкли: запись $x \leqslant y$ означает ни что иное, как $x, y \in (\leqslant)$.

В случае, когда $R \subseteq A \times A$, говорят, что отношение R ***задано на множестве*** A. Сосредоточимся на свойствах таких отношений.

Определение 5. Отношение $R \subseteq A \times A$

- ***рефлексивно***, если $\forall a \in A : aRa$;
- ***симметрично***, если $\forall a, b \in A : aRb \Rightarrow bRa$;
- ***транзитивно***, если $\forall a, b, c \in A : (aRb) \wedge (bRc) \Rightarrow aRc$.

Пример 18. Проверим эти свойства для отношения $(\leqslant)$ на множестве $\mathbb{R}$. Отношение $(\leqslant)$ рефлексивно: действительно, для любого числа a справедливо $a \leqslant a$. Отношение $(\leqslant)$ не симметрично: действительно, $1 \leqslant 2$, но $2 \not\leqslant 1$. Отношение $(\leqslant)$ транзитивно: какие бы три числа a, b и c мы не взяли, получим, что если $a \leqslant b$ и $b \leqslant c$, то $a \leqslant c$.

Упражнение 6. Проверьте каждое свойства для отношений $(>), (=), (\neq)$ на множестве $\mathbb{R}$.

10.3 Отношения эквивалентности

Определение 6. Рефлексивное, симметричное и транзитивное отношение называют ***отношением эквивалентности***.

Сделав упражнение 18 вы установили, что среди отношений $(\leqslant), (>), (=), (\neq)$ отношением эквивалентности является только равенство. Отсюда и следует смысл этого понятия: отношения эквивалентности определяют какие объекты считаются одинаковыми, а какие нет.

Пример 19. Определим отношение $\sim$ на множестве $\mathbb{Z} \times \mathbb{N}_1$ следующим образом: $(z, n) \sim (z', n')$, если и только если $\dfrac{z}{n} = \dfrac{z'}{n'}$. Если читатель не сообразил, что за отношение мы построили, то сообщим, что мы построили отношение равенства на множестве обыкновенных дробей. Оставляем читателю проверить, что это отношение рефлексивно, симметрично и транзитивно; самостоятельную проверку мы оставляем и для последующих примеров.

Пример 20. Следующий пример отношения эквивалентности — отношение равенства треугольников на плоскости. До середины XX века вместо слова «равенство» в учебниках по геометрии использовали термин «конгруэнтность».

Пример 21. Приведём ещё один пример отношения эквивалентности из геометрии: отношение подобия треугольников.

Пример 22. Перейдём теперь к математическому анализу. Напомним, что первообразной функции f, называется такая функция F, что $F'(x) = f(x)$ для всех $x \in \mathrm{Dom}(f)$. Отношение быть первообразной одной и той же функции f является отношением эквивалентности. Кстати, из курса анализа известно, что любые две первообразные одной и той же функции, отличаются друг от друга на константу.

Пример 23. Отношение равенств на булевых функциях, введённое нами на первой лекции, является отношением эквивалентности.

Интуитивно ясно, что если отношение $\sim \subseteq A \times A$ — отношение эквивалентности, то все элементы множества A можно разбить на подмножества попарно эквивалентных между собой объектов. В одинаковые подмножества попадают дроби, задающие одно и то же рациональное число, равные или подобные треугольники (в зависимости от отношения), а также первообразные одинаковых функций.

Эта интуиция отражает основную теорему об отношениях эквивалентности. Чтобы её сформулировать, формализуем сначала понятие класса эквивалентности. Пусть $\sim \subseteq A \times A$ — отношение эквивалентности. Определим ***класс эквивалентности*** $[a]$ как множество всех таких элементов множества A, которые эквивалентны элементу a:

$$[a] = \{x \mid x \in A,\, x \sim a\}.$$

Теорема 7. *Классы эквивалентности $[a]$ и $[b]$ (по отношению эквивалентности $\sim$) либо не пересекаются*[1]*, либо совпадают. Множество A разбивается в объединение классов эквивалентности.*

Доказательство. Ясно, что

$$A = \bigcup_{a \in A} [a],$$

поскольку каждый класс $[a]$ содержит элемент a в силу рефлексивности. Докажем теперь первую часть теоремы от противного.

Допустим $x \in [a] \cap [b]$ и $[a] \neq [b]$. Раз $x \in [a]$ и $x \in [b]$, то $x \sim a$ и $x \sim b$. В силу симметричности получаем, что $a \sim x$, а по транзитивности получаем, что

[1]Напомним, что множества A и B не пересекаются, если $A \cap B = \varnothing$.

$a \sim b$, раз $a \sim x \sim b$ (эта запись значит «$a \sim x$ и $x \sim b$»). Значит $a \in [b]$ и из симметричности и транзитивности получаем, что каждый элемент y из класса $[a]$ также принадлежит классу $[b]$:

$$y \in [a] \Rightarrow y \sim a \Rightarrow a \sim y \Rightarrow b \sim a \sim y \Rightarrow b \sim y \Rightarrow y \sim b.$$

То есть мы показали, что $[a] \subseteq [b]$. Симметричные рассуждения показывают, что $[b] \subseteq [a]$, а значит классы $[a]$ и $[b]$ совпадают, если пересекаются. □

Эта теорема объясняет, что отношение эквивалентности разбивает множество A на подмножества. Формализуем это утверждение. ***Семейством*** (множеств) называется множества, элементами которого являются множество.

Определение 7. ***Разбиением*** множества A на подмножества называется семейство $\mathcal{F} \subseteq 2^A$, для которого справедливы следующие условия

- $\varnothing \notin \mathcal{F}$;
- $A = \bigcup\limits_{B \in \mathcal{F}} B$;
- $\forall B, B' \in \mathcal{F} : B \neq B' \Rightarrow B \cap B' = \varnothing$.

То есть, разбиение множества A — это семейство непустых попарно непересекающихся его подмножеств, дающих в объединении A. Множества $B \in \mathcal{F}$ называют ***блоками*** разбиения $\mathcal{F}$.

Итак, классы эквивалентности образуют разбиение множества A. С другой стороны, каждому разбиению множества A соответствует отношение эквивалентности «элементы принадлежат одному блоку разбиения». Формально разбиению $\mathcal{F}$ ставится в соответствие отношение

$$\sim_{\mathcal{F}} = \{(x, y) \mid \exists B \in \mathcal{F} : x, y \in B\}.$$

Следствие 3. *Между разбиениями множества A и бинарными отношениями на A есть биекция, которая ставит в соответствие разбиению $\mathcal{F}$ отношение эквивалентности $\sim_{\mathcal{F}}$, классы которого являются блоками разбиения $\mathcal{F}$.*

Доказательство следствия оставим читателю в качестве упражнения.

Отношения эквивалентности полезны и в математике и в программировании. Если вы реализуете на C++ класс рациональных чисел через обыкновенные дроби, элементы множества $\mathbb{Z} \times \mathbb{N}_1$, то пары $(1, 2)$ и $(2, 4)$ задают одно и то же число: $1/2 = 2/4$; значит, на самом деле вам нужно либо реализовать рациональные числа, перейдя к классам эквивалентности по описанному отношению, и, например, каждый раз хранить в классе несократимую дробь, либо реализовать операцию равенства через проверку на эквивалентность.

В случае математики, отношения эквивалентности позволяют определить многие понятия и использовать основную теорему для доказательства соответствующих свойств. Так, неопределённый интеграл $\int f(x)dx$ формально является классом эквивалентности, потому и пишут

$$\int f(x)dx = F(x) + c.$$

В качестве применения основной теоремы введём отношение достижимости на множестве вершин неориентированного графа G. Вершина u достижима из v, если в G есть путь из v в u. Проверьте, что это отношение является отношением эквивалентности. Какие же у него классы? Подмножество $U \subseteq V(G)$ является классом эквивалентности по отношению достижимости, если для любой пары вершин $u, v \in U$ вершина u достижима из v и кроме того, нет других вершин в $V \setminus U$, достижимых из некоторой вершины $u \in U$. Таким образом, индуцированный подграф $G[U]$ является компонентой связности в графе G! Итак, мы доказали давно обещанный факт, который вытекает из основной теоремы об отношениях эквивалентности (теоремы 7).

Теорема 8. *Компоненты связности (простого неориентированного) графа либо не содержат общих вершин, либо совпадают. Любой граф является объединением своих компонент связности.*

Приведём ещё один пример отношения эквивалентности, с которым мы уже сталкивались в комбинаторике.

Пример 24. Зафиксируем множество $[n] = \{1, 2, \ldots, n\}$ и число k и обозначим через A множество слов длины k над алфавитом $[n]$, в которых все символы разные. Введём следующее бинарное отношение: слова $x_1x_2\ldots x_k$ и $y_1y_2\ldots y_k$ находятся в отношение $\sim$ тогда и только тогда, когда множества $\{x_1, x_2, \ldots, x_k\}$ и $\{y_1, y_2, \ldots, y_k\}$ совпадают:

$$x_1x_2\ldots x_k \sim y_1y_2\ldots y_k \iff \{x_1, x_2, \ldots, x_k\} = \{y_1, y_2, \ldots, y_k\}.$$

Это отношение является отношением эквивалентности и его классами являются подмножества k-элементные подмножества множества $[n]$. Мы фактически использовали это отношение в примере 16 для вывода формулы для чисел сочетания. Основной факт, который мы использовали: все классы эквивалентности имеют одинаковый размер $k!$. Отсюда следует, что количество классов равно $\dfrac{|A|}{k!}$.

10.4 Операции с бинарными отношениями

Поскольку бинарные отношения формально являются множествами, к ним применимы все теоретико-множественные операции. Обратим внимание, что операции удобно вычислять, в случае, если бинарные отношения заданы матрицами: если $R = P \cap Q$, то $xRy \iff (xPy) \wedge (xQy)$, то есть для вычисления матрицы отношения R нужно взять поэлементную конъюнкцию матриц отношений P и Q. Здесь мы использовали уже хорошо изученную нами связь между алгеброй множеств и алгеброй логики.

Бдительный читатель уже сообразил, что бинарные отношения ничем не отличаются от бинарных предикатов: предикат $R(x, y) = 1$ тогда и только тогда, когда выполняется отношение xRy. Как всегда, в разных ветвях математики возникают разные определения и обозначения для эквивалентных объектов, и эти расхождения со временем изжить очень тяжело.

Помимо теоретико-множественных операций, изучим ряд естественных операций для бинарных отношений.

Операция обращения (транспонирования)

Обратным отношением к отношению $R \subseteq A \times B$ называют отношение

$$R^{-1} = \{(y, x) \mid xRy\} \subseteq B \times A.$$

Операция обращения отношений известна также как операция транспонирования, поскольку в случае отношений между конечными множествами, обратное отношение задаётся транспонированной матрицей исходного.

Упражнение 7. Докажите, что отношение $R \subseteq A \times A$ симметрично тогда и только тогда, когда $R = R^{-1}$. Выразите свойство симметричности на языке матриц.

В случае, если отношение f функционально (является функцией) и обратное к нему отношение является функцией, то отношение f^{-1} реализует обратную функцию. В этом случае функция f называется ***обратимой***.

Упражнение 8. Докажите, что отношение R^{-1} функционально тогда и только тогда, когда отношение R инъективно.

Напомним, что в случае функций запись f^{-1} означает полный прообраз. И эта запись осмыслена с точки зрения бинарных отношений. Любому бинарному отношению $R \subseteq A \times B$ соответствует функция $f_R : 2^A \to 2^B$, такая что для любого подмножества $X \subseteq A$ справедливо

$$f_R(X) = \{y \mid \exists x \in X : xRy\}.$$

В случае когда R является функцией, отображение f_R возвращает образ множества X, поэтому для сокращения и пишут $f(X)$ и даже $R(X)$. Отображение $f_{R^{-1}}$ (для обратного к R отношения) вычисляет полный прообраз множества $Y \subseteq B$.

Операция композиции и связь с базами данных

Первым способом описания бинарных отношений, который мы изучили, были таблицы. Таблица из примера напоминает упрощённую версию таблицы базы данных: в реальной базе данных таблицы имеют много столбцов и значение в каждом из них не обязательно 0 или 1. Тем не менее, реляционные базы данных берут своё название от отношений (relations) правда произвольной арности. В общем случае, k-арное отношение определяется как подмножество декартова произведения k множеств:

$$R \subseteq A_1 \times A_2 \times \ldots \times A_k.$$

Для простоты мы считаем операцию декартова произведения ассоциативной, и считаем, что элементы множества $A_1 \times A_2 \times \ldots \times A_k$ — это (упорядоченные) наборы или как их ещё называют ***кортежи***:

$$A_1 \times A_2 \times \ldots \times A_k = \{(a_1, a_2, \ldots, a_k) \mid a_i \in A_i\}.$$

Заметим, что мы начинали изучение курса с отношений арности 1 — это просто множества. В работе с базами данных важную роль играет операция композиции, которую мы определим только для бинарных отношений, но сути это не меняет:

на k-арные отношения можно смотреть как на бинарные, считая, что $R \subseteq A_1 \times B$, где $B = A_2 \times A_3 \times \ldots \times A_k$.

Перед формальным определением композиции отношений, начнём с примера и проведём его на языке графов.

В университете есть множество студентов X, распределённых по множеству факультетов Y. Это распределение задаёт отношение $P \subseteq X \times Y$. У университета также есть множество зданий Z, при этом разные факультеты проводят свои занятия в разных зданиях: факультет y находится в отношении Q со зданием z, если и только если у какой-то из групп есть занятия в здании z. Для повышения мер безопасности администрация факультета решила предоставить студентам допуск только в те здания, в которых у есть занятия у их факультета. Для этих целей, администрация изобразила отношения P и Q с помощью трёхдольного графа (рис. 10.3).

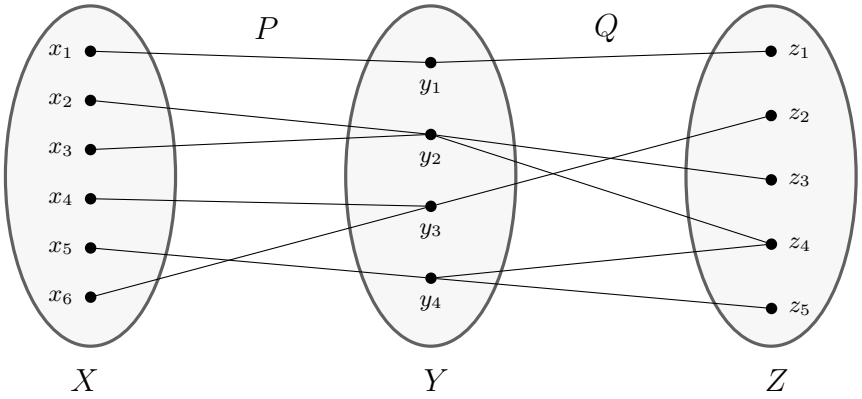


Рис. 10.3. Композиция бинарных отношений

Ясно, что студенту x_i нужно дать доступ к зданию z_j, если и только если из x_i можно добраться до z_j через какую-то вершину множества Y. Обозначим через R отношение доступа студента к зданию. Ясно, что $x_2 R z_4$, потому что $x_2 P y_2$ и $y_2 Q z_4$. Отношение R и будет результатом композиции отношений P и Q, которую мы готовы теперь определить формально.

$$Q \circ P = \{(x, z) \mid \exists y \in Y : xPy \wedge yQz\}.$$

Обратите внимание на порядок операндов композиции. Он не совпадает с порядком отношений на картинке, но совпадает с порядком операндов при композиции функций: $(f \circ g)(x) = f(g(x))$. Мы следуем данному соглашению о порядке операндов, чтобы не было недоразумений при стандартной композиции функций.

Лекция 11

Ориентированные графы и отношения порядка

План:

1. Определение ориентированного графа. Исходящие и входящие степени — аналог формулы суммы степеней для неориентированного графа.
2. Компоненты сильной связности
3. **Т.:** Следующие условия для ориентированного графа равносильны:
   - Каждая компонента сильной связности тривиальна (состоит из одной вершины)
   - Граф ациклический
   - Вершины графа можно занумеровать так, что рёбра идут только от вершин с меньшим номером к вершинам с большим номером.
4. Отношения частичного порядка
   - Примеры отношений порядка (покоординатный порядок)
   - Линейный порядок
   - Отношение непосредственного следования и его граф (диаграмма Хассе)
   - Покоординатный порядок
   - Булев куб — двоичные слова, упорядоченные покоординатно

Неформально, ориентированные графы соответствуют перекрёсткам и дорогам с односторонним движениям; если между улицами u и v двустороннее движение, то есть как ребро из u в v, так и ребро из v в u. Формально, ***ориентированный***

граф задан парой (V, E), где множество вершин V как и раньше произвольное (по-умолчанию, конечное и непустое) множество, а множество $E \subseteq V \times V$ — состоит из упорядоченных пар вершин. По-умолчанию, мы считаем, что в каждой паре вершины различны:

$$E \subseteq \{(u, v) \mid u, v \in V,\ u \neq v\}.$$

Рёбра вида (u, u) называются ***петлями*** и они возникают естественным образом при описании бинарных отношений с помощью ориентированных графов. Мы будем отдельно оговаривать использования петель: говорить «***ориентированный граф, возможно с петлями***». Также часто рассматривают ориентированные графы, в которых из одну вершину в другую идёт более одного ребра; такие рёбра называют параллельными (или кратными), но нам они не пригодятся.

Пример ориентированного графа приведён на рис. 11.1.

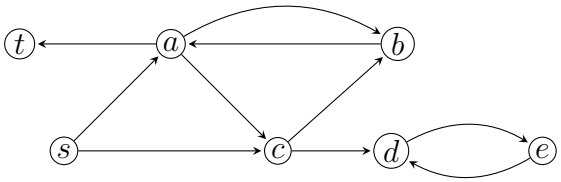


Рис. 11.1. Ориентированный граф G.

11.1 Базовые понятия для ориентированных графов

Определения, введённые нами для неориентированных графов, с поправками переносятся на ориентированные. ***Исходящей степенью*** вершины $d_+(u)$ называют число рёбер, исходящих из вершины u, ***входящей степенью*** $d_-(u)$ — число рёбер, входящих в u. Запомнить знак можно по по правилу «ток идёт от плюса к минусу». Вершины входящей степени 0 называют ***источниками***, к таким относится вершина s $(d_-(s) = 0)$, а вершины с нулевой исходящей степенью называют ***стоками***: $d_+(t) = 0$. Источники и стоки часто обозначают соответственно через s и t, от слов source и target, хотя стоки на английском и называются sink.

Теорема о сумме степеней вершин при переносе на ориентированные графы тривиализуется.

Утверждение 10. $\sum_{u \in V} d_+(u) = \sum_{u \in V} d_-(u) = |E|$.

Подграфы, пути и циклы определяются аналогично неориентированным графам, только вместо неориентированных путей и циклов используют ориентированные.

В случае ориентированных графов цикл может быть и на двух вершинах: примером такого цикла служит подграф графа G, индуцированный вершинами d и e (на рис. 11.1).

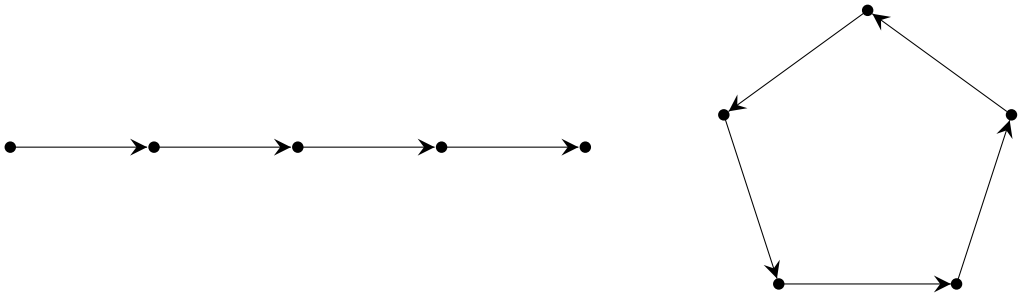


Рис. 11.2. Ориентированные путь и цикл

В случае ориентированных графов проще иметь дело с маршрутами, определение которых не меняется при переходе от неориентированных графов. ***Маршрутом*** в ориентированном графе называется последовательность вершин $v_0, v_1, \ldots, v_n$, такая что $n \geqslant 0$ и $(v_i, v_{i+1}) \in E(G)$ для $0 \leqslant i \leqslant n-1$. Обратим внимание, что второе условие применимо только в случае, когда в маршруте больше одной вершины, а потому любая последовательность из единственной вершины считается маршрутом. ***Длина маршрута*** — это число рёбер, соединяющих вершины маршрута; оно совпадает с n. Маршрут называется ***замкнутым***, если $v_0 = v_n$.

Следующая лемма связывает маршруты, пути и циклы. Её доказательство мы оставляем читателю в качестве упражнение (оно почти повторяет рассуждения для аналогичной леммы для неориентированных графов).

Лемма 6. *Между двумя вершинами ориентированного графа существует путь тогда и только тогда, когда между ними существует маршрут. В графе есть цикл тогда и только тогда, когда в нём есть замкнутый маршрут положительной длины.*

Компоненты сильной связности

Следующий термин на очереди — компоненты связности, но его так просто не перенесёшь. Из рис. 11.1 ясно, что из вершины s есть путь в вершину t, а пути обратно нет. Поэтому вместо компонент связности, в ориентированных графах рассматривают компоненты сильной связности.

Для определения компонент сильной связности нам потребуется отношение двусторонней достижимости. Вершина v ***достижима*** из u, если существует маршрут из u в v; отношение достижимости между вершинами обозначим $u \rightsquigarrow v$. Определим отношение двусторонней достижимости $u \leftrightsquigarrow v = (u \rightsquigarrow v) \wedge (v \rightsquigarrow u)$.

Утверждение 11. *Отношение двусторонней достижимости* ($\leftrightsquigarrow$) *— отношение эквивалентности.*

Определение 8. ***Компонента сильной связности*** ориентированного графа — класс эквивалентности по отношению двусторонней достижимости. То есть множество $U \subseteq V$ — компонента сильной связности, если любые две вершины множества U достижимы друг из друга и в U нельзя добавить ещё вершины с сохранением этого свойства (множество U — максимальное по включению).

Из основной теоремы об отношениях эквивалентности получаем, что компоненты сильной связности не пересекаются или совпадают.

Пример 25. Найдём компоненты сильной связности графа G на рис. 11.1. Вершины d и e попарно достижимы друг из друга (т. к. лежат на цикле) и из них не достижимы другие вершины — они образуют компоненту сильной связности. Вершины a и b тоже лежат на цикле, но они не образуют компоненту сильной связности, поскольку к ним ещё можно добавить вершину c; вторая компонента: множество $\{a, b, c\}$. Вершины s и t образуют отдельные компоненты сильной связности: каждая из них достижима из себя по определению маршрута, а других вершин из которых достижима s (или которые достижимы из t) нет.

Каждому ориентированному графу G можно поставить в соответствие граф, вершинами которого являются компоненты сильной связности G, а ребро ведёт из вершины-компоненты U в U', если в графе G есть ребро $u \to u'$, $u \in U$, $u' \in U'$. Такой граф называют ***конденсатом*** графа G; обратим внимание, что кратных рёбер в конденсате нет.

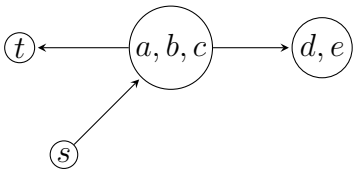


Рис. 11.3. Граф конденсат графа G.

Легко видеть, что в конденсате графа G (рис. 11.3) нет циклов. Так получилось случайно или это справедливо для любого конденсата? Давайте исследуем свойства компонент сильной связности, чтобы ответить на этот вопрос.

Утверждение 12. *Вершины u и v лежат в одной компоненте сильной связности тогда и только тогда, когда они лежат на замкнутом маршруте.*

Доказательство. Действительно, условие двусторонней достижимости влечёт существование маршрутов из u в v и из v в u, склеив их получим замкнутый маршрут. С другой стороны, замкнутый маршрут, содержащий вершины u и v очевидно влечёт двустороннюю достижимость. □

Из этого свойства ясно, что в конденсате не может быть циклов: если бы они были, то компоненты U и U' содержали бы вершины из одной компоненты сильной связности, что противоречит свойствам классов эквивалентности (они либо не пересекаются, либо совпадают).

11.2 Ациклические графы

Ориентированный граф называется ***ациклическим***, если в нём нет циклов, или, что равносильно по лемме 6, в нём нет замкнутых маршрутов положительной длины. Следующая теорема устанавливает равносильные условия ацикличности графа.

Теорема 9. *Для ориентированного графа $G(V, E)$ равносильны следующие условия:*

(1) граф G ациклический;

(2) каждая компонента сильной связности графа G ***тривиальна****, т. е. состоит из единственной вершины;*

(3) вершины графа G можно занумеровать числами от 1 до $|V|$ так, что рёбра G идут только от вершин с меньшим номером в вершины с большим номером.

Доказательство. Равносильность условий (1) и (2) вытекает из утверждения 12: в графе есть нетривиальная компонента сильной связности тогда и только тогда когда в нём есть замкнутый маршрут длины 2 или больше, а это равносильно существованию цикла.

Из условия 3 очевидно следует условие 1. Действительно, если такая нумерация существует, то если был бы цикл, то в нём вершина с большим номером соединялась бы ребром с меньшим номером.

Для завершения доказательства теоремы осталось доказать импликацию (1) $\Rightarrow$ (3). Докажем для этого вспомогательную лемму.

Лемма 7. *В ориентированном ациклическом графе G есть сток.*

Доказательство. Возьмём самый длинный ориентированный путь $v_1 \to v_2 \to \ldots \to v_n$ в графе G. Такой существует, потому что множество путей конечно (вершины в пути повторяться не могут). Докажем от противного, что вершина v_n является стоком ($d_+(v_n) = 0$). Если в G есть ребро $v_n \to u$ и $u \neq v_i$, то путь можно сделать длиннее, добавив к нему u; если же $u = v_i$, то в графе G есть цикл (петель в G по определению быть не может). Оба случая приводят нас к противоречию. □

Докажем с помощью леммы импликацию (1) $\Rightarrow$ (3) индукцией по числу вершин. База: для $|V| = 1$ очевидна: в графе из одной вершины нет рёбер, поэтому занумеровав единственную вершину единицей получим корректную нумерацию. Шаг: пусть утверждение верно для любого графа на n вершинах; согласно лемме в графе G на $(n+1)$-ой вершине существует сток — занумеруем его числом $n+1$ и рассмотрим граф G', получающийся из G удалением этого стока. По предложению индукции вершины G' можно занумеровать корректно; перенеся эту нумерацию на G получим также корректную нумерацию: поскольку из $(n+1)$-ой вершины не идёт ни одного ребра, испортить нумерацию она не может, а для всех остальных рёбер нумерация корректна. □

Нумерация вершин, удовлетворяющая условию (3) называется ***топологической сортировкой***. Этот приём очень полезен на практике. Представьте, что граф G — граф зависимости пакетов (программ) в операционной системе: ребро $u \to v$ означает, что пакет u зависит от пакета v и чтобы установить новую версию пакета u нужно сначала установить новую версию пакета v. В случае, если в графе G есть циклы, обновить операционную систему не получится: многим пользователям GNU/Linux до боли знакома эта проблема (часто она приводит к переустановке всей операционной системы). Если же граф ациклический, то для корректного обновления нужно выполнить топологическую сортировку и обновлять пакеты

в порядке убывания номеров (чем выше приоритет, тем раньше нужно обновить пакет).

11.3 Ориентированные графы и бинарные отношения

Бинарное отношение R на конечном множестве V может быть задано 0-1-матрицей, которая в свою очередь является матрицей смежности для ориентированного графа (возможно с петлями). Другими словами, пары входящие в отношение R являются рёбрами графа $G(V, R)$.

Как изученные нами на прошлой лекции свойства выражаются в терминах графов? Опишем связь для свойств из раздела 10.2 прошлой лекции, а также введём здесь нужные нам свойства для этой лекции и сразу исследуем их связь с графами. Бинарным отношениям на бесконечных множествах соответствуют бесконечные графы; используемые нами определения для них не отличаются от определений для конечных графов.

Отношение рефлексивно, если в каждой вершине графа есть петля (диагональ матрицы состоит из единиц). В случае рефлексивных бинарных отношений часто удобно считать, что петель нет, помня, что отношение рефлексивно; петли полезны, чтобы выделить отношения, которые не рефлексивны и не антирефлексивны: отношение $R \subseteq V \times V$ называется ***антирефлексивным***, если не содержит ни одной пары (v, v) (диагональ матрицы состоит из нулей). В графе антирефлексивного отношения петель нет.

Отношение R симметрично, если в соответствующем графе ребро из u в v есть тогда и только тогда, когда есть ребро из v в u (матрица симметрична). Симметричным бинарным отношениям соответствуют неориентированные графы: вместо двух ориентированных рёбер между u и v проводят одно неориентированное.

Отношение R является ***антисимметричным***, если из uRv и vRu следует, что $u = v$; формально

$$\forall u, v \in V : (uRv) \wedge (vRu) \Rightarrow (u = v).$$

То есть из пар (u, v) и (v, u) в отношении может быть не более одной, за исключением петель. На языке ориентированных графов это означает, что если в графе существует ребро из u в v, то в графе нет ребра из v в u.

Отношение R транзитивно, если в соответствующем графе существование маршрута из вершины u в вершину v влечёт существование ребра из u в v. Формальное определение транзитивности утверждает только, что из наличия в графе рёбер $u \to w$ и $w \to v$ следует наличие ребра $u \to v$, но это утверждение легко обобщается по индукции до заявленного в начале абзаца.

11.4 Отношения порядка

В рамках этой лекции мы сосредоточимся на отношениях, которые транзитивны и антисимметричны и либо рефлексивны, либо антирефлексивны. Такие отношения

называют ***отношениями (частичного) порядка*** или ***(частичными) порядками***. Отношения порядка знакомы читателю со школьной скамьи: таковыми являются отношения $(\leqslant)$ и $(<)$ на множествах $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$. Эти символы традиционно используют для отношений порядка Рефлексивные отношения порядка традиционно обозначают символом $\leqslant$, быть может с индексом, такие порядки называют ***нестрогими***, антирефлексивные отношения порядка обозначают символом $<$, их называют ***строгими***. Ясно, что каждому отношению нестрогого порядка $\leqslant_P$ на некотором множестве A однозначно соответствует строгое отношение порядка $<_P$, которое получается из P удалением всех пар (a, a), и это соответствие взаимно однозначно (между строгими и нестрогими порядками определена биекция). Поэтому мы будем переходить в рассуждениях от нестрогого порядка $\leqslant_P$ к соответствующему строгому $<_P$ без дополнительных пояснений.

Разберёмся, как отношения порядка на конечных множествах связаны с ориентированными графами. Построим граф для отношения порядка $<$ на множестве $\{0, 1, 2, 3, 4\}$:

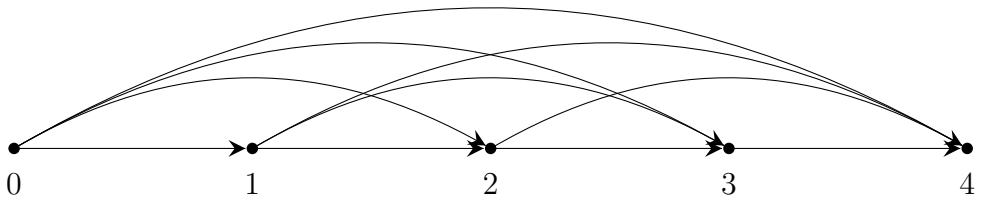


Рис. 11.4. Граф $G(\{0, 1, 2, 3, 4\}, (<))$

Как видно из рисунка граф получился ациклическим. Докажем справедливость этого наблюдения для общего случая.

Утверждение 13. *Отношению строго порядка $(<_P)$ соответствует ориентированный ациклический граф.*

Доказательство. Предположим противное: в графе отношения $<_P$ нашёлся замкнутый маршрут: $v_1, v_2, \ldots, v_{n-1}, v_n = v_1$. Поскольку из v_1 ведёт ребро в v_2, а из v_2 в v_3 получаем по транзитивности отношения $<_P$, что из v_1 ведёт ребро в v_3. Продолжая по индукции получаем, что из v_1 ведёт ребро в $v_n = v_1$, то есть в графе G есть петля. Но отношение $<_P$ антирефлексивно, что означает, что в его графе нет петель. □

Заметим, что не каждому ориентированному ациклическому графу соответствует отношение порядка: отношение, соответствующее графу может не быть транзитивным.

Граф 11.4 выглядит громоздким и избыточным: слишком много рёбер добавляется по транзитивности. Поэтому при описании отношения порядка графом использует не граф для самого отношения порядка $<_P \subseteq V \times V$ (или $\leqslant_P$), а граф для ***отношения непосредственного следования*** $\prec_P$, которое определяется согласно формуле:

$$(\prec_P) = \{(x, y) \mid (x <_P y) \wedge (\neg \exists z \in V : (x <_P z) \wedge (z <_P y))\}.$$

Другими словами, отношение $x \prec_P y$ означает, что $x <_P y$ и между ними нет элементов. Отношение непосредственного следования определено для любого порядка на конечном множестве, но может быть не определено (формально оно пусто) для отношения порядка на бесконечном множестве, таком как множество действительных чисел.

Граф отношения непосредственного следования для отношения $<$ на множестве $\{0, 1, 2, 3, 4\}$ совпадает с ориентированным путём. Граф $G(V, \prec_P)$ отношения непосредственного следования называют ***диаграммой Хассе***.

Примеры, базовые определения и свойства

Отношение порядка, в котором любые два элемента сравнимы называется ***линейным***. Диаграмма Хассе для линейного порядка на конечным множеством является графом путём. Слово «частичный» перед порядком добавляют, чтобы подчеркнуть, что порядок не обязательно линейный.

Следующий естественный пример — покоординатный порядок проиллюстрирован диаграммой Хассе на рис. 11.5. Рисунок годится как для порядка на множестве $\mathbb{N}_0 \times \mathbb{N}_0$ (если представить, что у картинки есть продолжение), так и для порядка на множестве $\{0, 1, \ldots, 5\} \times \{0, 1, \ldots, 4\}$.

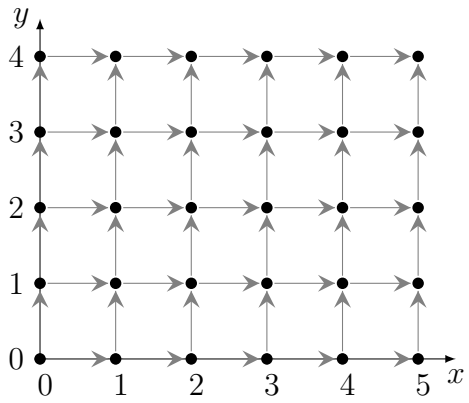


Рис. 11.5. Покоординатный порядок

Для формального определения покоординатного порядка удобно сначала ввести определения (декартова) произведения на порядках.

Определение 9. Пусть $\leqslant_P$ и $\leqslant_Q$ отношения порядка на множествах A и B соответственно. Определим их ***произведение*** как отношение $(\leqslant_{P\times Q})$ на множестве $A \times B$:

$$(a, b) \leqslant_{P\times Q} (a', b') \iff \begin{cases} a \leqslant_P a'; \\ b \leqslant_Q b'. \end{cases}$$

Таким образом, рис. 11.5 иллюстрирует произведение двух линейных порядков $\leqslant$ на множестве $\mathbb{N}_0$ или произведение линейных порядков на множествах $\{0, 1, \ldots, 5\}$ и $\{0, 1, \ldots, 4\}$.

Определение 10. Зафиксируем отношение порядка $\leqslant_P$ на множестве A. Элемент $x \in A$ называется

- ***максимальным***, если не существует элемента $a \in A$, такого что $x < a$;
- ***наибольшим***, если для любого элемента $a \in A$ справедливо $a \leqslant x$.
- ***минимальным***, если не существует элемента $a \in A$, такого что $a < x$;
- ***наименьшим***, если для любого элемента $a \in A$ справедливо $x \leqslant a$.

Из определения ясно, что наибольший элемент порядка является максимальным; обратное вообще говоря неверно. Приведём контрпример для минимального и наименьшего элемента. Удалим из порядка $\leqslant_{\mathbb{N}_0 \times \mathbb{N}_0}$ точку $(0, 0)$. Получим порядок $\leqslant_P$, в котором элементы $(1, 0)$ и $(0, 1)$ являются минимальными, но между собой они несравнимы, поэтому наименьшего элемента в этом порядке нет.

Перейдём к следующему примеру. Напомним, что 2^U обозначает множество всех подмножеств множества U. Для любого множества U отношение $\subseteq$ на множестве U является нестрогим отношением порядка, а ему соответствует строгое отношение $\subsetneq$. Пусть $U = \{1, 2\}$; тогда получаем

$$\varnothing \subseteq \{1\}, \quad \varnothing \subseteq \{2\}, \quad \{1\} \subseteq \{1, 2\}, \quad \{2\} \subseteq \{1, 2\};$$

перечисленные пары задают диаграмму Хассе для включения.

Зафиксируем неориентированный граф G. Отношение $\subseteq$ «быть подграфом» является отношением порядка на множестве подграфов графа G.

Ясно, что если порядок $\leqslant_P$ определён на множестве A, то его можно сузить до любого подмножества X множества A. Формально сужение устроено так: $(\leqslant_P) \cap X \times X$.

Сужение порядка позволяет перенести понятие максимального и минимального элемента на подмножества. Так, сузив отношение $\subseteq$ на связные подграфы графа G получим, что компоненты связности являются максимальными подграфами, которые связны; а сузив отношение $\subseteq$ на множество клик получим, что максимальные элементы этого отношения являются максимальными кликами.

На рис. 11.6 изображён граф, известный как ориентированный булев куб. Простой булев куб получается из графа на рисунке стиранием ориентации. Проверьте, что булев куб является диаграммой Хассе для произведениея линейного порядка на множестве $\{0, 1\}$ на себя три раза. Дадим общее определение ориентированного булева куба. Множеством вершин ориентированного булева куба B_n являются двоичные слова; ребро идёт от слова u к слову v, если они отличаются ровно в одной позиции и в слове v на этой позиции стоит 1.

11.5 Связь между теоремой об ациклических графах и порядках

Отношения порядка на конечных множествах обладают следующим хорошим свойством. Если отношение порядка $\leqslant_P$ (на конечном множестве A) не является линейным, то его можно продолжить до линейного. Это означает, что можно

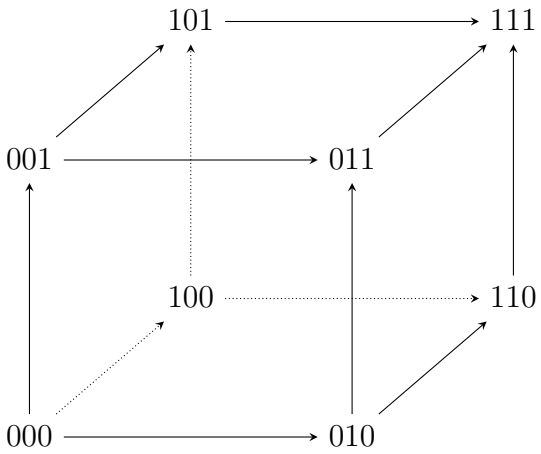


Рис. 11.6. Булев куб

добавить в отношение $\leqslant_P$ новые пары (x, y) не изменив при этом старые пары так, что все элементы множества A будут сравнимы между собой в получившемся порядке.

Это свойство напрямую вытекает из теоремы 9: отношению $\leqslant_P$ соответствует ациклический граф. Сделаем его топологическую сортировку и добавим всевозможные рёбра от вершин с меньшим номером к вершинам с бóльшим номером. При этом ни одно соотношение в порядке $\leqslant_P$ не будет нарушено, а все несравнимые ранее элементы станут сравнимыми.

Лекция 12

Булевы функции

План:

1. Алгоритм построения ДНФ (и КНФ) по таблице истинности
2. Определение булевых схем, реализующих булевы функции, через последовательности присваиваний и графов (стандартный базис).
3. Формулы — схемы специального вида
4. Общее определение схем (для произвольного базиса)
5. Базис — полный базис
6. Монотонные функции
   - неполнота монотонного базиса $\{\wedge, \vee\}$
   - связь с множествами (монотонность по включению)
   - раскраска булева куба
   - монотонных булевых функций от $2n$ переменных не меньше чем $2^{\frac{2^{2n}}{2n+1}}$.
7. Многочлены Жегалкина
   - Базис Жегалкина, его полнота
   - Многочлены Жегалкина в стандартном виде
   - Биекция между многочленами Жегалкина в стандартном виде и булевыми функциями

8*. Классы Поста. Формулировка теоремы Поста и план её доказательства

Мы возвращаемся к изучению булевых функций после изучения других тем. В этой лекции нам пригодится весь ранее изученный материал; с его помощью мы не просто докажем важные факты о булевых функциях, но и увидим связь между разными темами.

На первой лекции мы изучили два способа задания булевой функции: таблицей истинности (и её сжатым описанием — вектором значени) и формулами. При этом у нас не было формального определения формулы. В этой лекции мы изучим ещё один способ задания булевой функции — булевы схемы, которые мы используем и для формального определения формул. Также мы познакомимся с вопросом полноты класса функций и узнаем, как по опирациям, которые могут быть использованы в формуле, определить можо ли задать с помощью таких формул любую булеву функцию.

12.1 Построение ДНФ

Насколько равноправны способы задания булевых функций, которые мы изучили? Ясно, что любую булеву функцию можно задать таблицей истинности, или, что то же самое, вектором значений.

Что можно сказать о формулах? Всё зависит от того какой набор логических связок (т. е. булевых функций) мы будем использовать для построения формул. Такие наборы называются ***базисами***. Мы покажем, что любую функцию алгебры логики можно выразить через формулы над базисом «и, или, не» $\{\wedge, \vee, \neg\}$ — ***стандартным базисом***.

Для начала нам понадобятся вспомогательные определения. ***Литералом*** называют переменную или отрицание переменной. Обозначим $x^1 = x$, а $x^0 = \overline{x}$. Тогда литерал имеет вид $l = x^\alpha$, где $\alpha = 0$ или $\alpha = 1$.

Конъюнктом называют конъюнкция литералов:

$$C = l_1 \wedge l_2 \wedge \ldots \wedge l_k.$$

Заметим, что всякий конъюнкт $x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_k^{\alpha_k}$ принимает значение 1 только на одном наборе значений переменных $x_1, x_2, \ldots, x_k$ и это набор $\alpha_1, \alpha_2, \ldots, \alpha_k$.

Для каждого булева вектора α определим функцию $F_\alpha(x_1, \ldots, x_n) = x_1^{\alpha^1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$. Из сказанного выше следует, что

$$F_\alpha(x_1, \ldots, x_n) = \begin{cases} 1, & \text{если } \forall i \; x_i = \alpha_i \\ 0, & \text{иначе} \end{cases}$$

Таким образом каждую, отличную от тождественно нулевой, булеву функцию можно представить как дизъюнкцию конъюнктов:

$$f(x_1, \ldots, x_n) = \bigvee_{\alpha \in f^{-1}(1)} F_\alpha(x_1, \ldots, x_n) = \bigvee_{\alpha \in f^{-1}(1)} \bigwedge_{i=1}^{n} x_i^{\alpha_i} \tag{1}$$

Формула, которая имеет вид дизъюнкции конъюнктов называется ***дизъюнктивной нормальной формой (ДНФ)***. Формула (1) является разложением произвольной булевой функции в ДНФ. Более того в ДНФ (1) в каждом конъюнкте

встречается каждая переменная — такие ДНФ называют ***ДНФ в стандартном виде***. Для тождественно нулевой функции легко построить ДНФ: $x_1 \wedge \overline{x_1}$.

Взяв отрицание от частей формулы (1) и применив закон Моргана получим

$$\neg f(x_1, \ldots, x_n) = \bigwedge_{\alpha \in f^{-1}(1)} \neg F_\alpha(x_1, \ldots, x_n) = \bigwedge_{\alpha \in f^{-1}(1)} \bigvee_{i=1}^{n} \neg x_i^{\alpha_i}$$

Обозначив $g(x_1, \ldots, x_n) = \neg f(x_1, \ldots, x_n)$ и заметив, что $\neg x_i^{\alpha_i} = x_i^{\overline{\alpha_i}}$, получим формулу разложения

$$g(x_1, \ldots, x_n) = \bigwedge_{\alpha \in g^{-1}(0)} \bigvee_{i=1}^{n} x_i^{\overline{\alpha_i}}. \tag{2}$$

Эта формула является конъюнкцией дизъюнктов (***дизъюнкт*** — дизъюнкция литералов), и называется формулой разложения в ***конъюнктивную нормальную форму***, которя определяется как конъюнкция дизъюнктов. Как и в случае с ДНФ (1), формула (2) является формулой разложения функции g в ***КНФ в стандартном виде***.

12.2 Булевы схемы

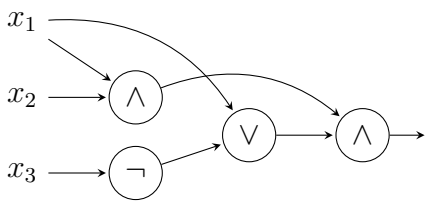


Рис. 12.1. Булева схема, заданная графом

Булевы схемы — не просто способ задания булевой функций, но и фундаментальная модель вычислений. На практике булевы схемы используются в микросхемах; как часто бывает на практике, чаще используют более сложные схемы, чем мы рассматриваем. Рис. 12.2 иллюстрирует схему, которая получает на вход значения переменных 1 или 0 — наличие тока или его отсутсвие. Далее каждый узел схемы реализует одну из булевых операций и передаёт вычисленное значение на следующий узел. На практике возникает естественная проблема: как определить, что значение на выходе схемы (самый правый узел) вычисленно окончательно? Вдруг какой-то сигнал ещё не успел пройти и значение после его прихода может измениться. Чтобы решить эту проблему используют часы (сигнал с определённой частотой), но мы в рамках курса остановимся на изучении схем как на рисунке.

Схема на рис. 12.2 задана графом. Формальное определение однако граф не использует, и изучив его мы увидим, что далеко не всякую схему можно описать графом; графом можно задать схему, в которой используются только операции $\wedge, \vee, \neg$ или другие коммутативные операции (у которых порядок аргументов неважен). Операции, которые могут использоваться в схеме, называют ***базисом***. Формально

базис — это множество булевых функций B. Для теоретических целей нам удобно считать, что базис может быть бесконечным. Читатель может заметить отличие от курса линейной алгебры, в которой базис — не просто множество векторов, а упорядоченное множество. Мы не будем уделять этому аспекту особое внимание, для наших целей достаточно уметь различать разные функции из базиса.

Булевой схемой в базисе B называется последовательность присваиваний, которая начинается со всех переменных используемых схемой $x_1, x_2, \ldots, x_n$ и продолжается ***присваиваниями*** $s_{n+1}, s_{n+2}, \ldots s_m$, где каждое присваивание имеет вид $s_i = g(s_{j_1}, s_{j_2}, \ldots, s_{j_k})$, где $j_1, \ldots, j_k < i$, $g \in B$; под $s_1, \ldots, s_n$ понимаются переменные $x_1, \ldots, x_n$. Булева схема задаёт булеву функцию $f(x_1, \ldots, x_n)$, значение которой совпадает со значением последнего присваивания s_m (после проделанных вычислений).

Так схема на рис. 12.2 задана последовательностью присваиваний

$$x_1,\ x_2,\ x_3,\ s_4 = x_1 \wedge x_2,\ s_5 = \neg x_2,\ s_6 = x_1 \vee s_5,\ s_7 = s_4 \wedge s_6 \tag{3}$$

Это схема в базисе $\{\wedge, \vee, \neg\}$, который называют ***стандартным***.

Заметим, что схеме на рис. 12.2 можно поставить в соответствие несколько последовательностей присваиваний; как записать по графу последовательность присваиваний и всегда ли это возможно? Это возможно если и только если граф ациклический: вершины графа можно занумеровать в порядке топологической сортировки[1] тогда и только тогда, когда граф ациклический. Таким образом, ориентированный граф, в вершинах которого записаны функции из базиса B и одна из вершин помечена как выход, задаёт булеву схему, если он ациклический, а все функции в вершинах ***симметрические***, т. е. перестановка битов в булевом векторе на входе не меняет значение функции. Коммутативная булева операция от двух аргументов является симметрической булевой функцией.

Вспомним, что на первой лекции мы определяли булеву формулу через вычисление по дереву. Это определение можно сделать формальным с помощью булевой схемы. ***Булева формула*** — это булева схема, в которой ни одно присваивание, кроме переменных, не используется дважды в правой части присваиваний. Проверьте, что последовательность присваиваний (3) задаёт булеву формулу. Булева формула имеет простое описание на языке графов (в случае, когда схему можно задать графом): схема, заданная графом, является формулой, если любая вершина кроме переменной имеет исходящую степень не больше 1. Это условие можно переформулировать и так, но с оговоркой: после удаления из графа вершин с переменными (вместе с рёбрами) и удалением ориентации получается дерево. Нужно оговорить, что мы предполагаем, что в схеме нет бесполезных узлов, — иначе, во-первых, может получиться не дерево, а лес — граф, компонентами связности которого являются деревья, а во вторых могут появиться и вершины с исходящей степенью больше 1, но при этом только одно ребро таких вершин будет вести «в сторону выхода». Скобки позволяют закодировать дерево, поэтому формулы, к которым мы привыкли просто кодируют деревья вычисления или если формальнее — схемы. Схемы эти не обязательно булевы: так, арифметические формулы задают арифметические схемы, которые определяются аналогично булевым, а в случае

[1] Напомним, что это порядок, в котором рёбра идут только от вершин с меньшим номером к вершинам с большими номерами

использования каких-то специальных функций (например, тригонометрических) можно использовать аналогичное определение схемы, в котором в базис могут входить произвольные (не обязательно булевы) функции.

Ясно, что любую булеву схему можно превратить в формулу, увеличив число присваиваний. Покажем это с помощью графов, дабы проиллюстрировать выше сказанное

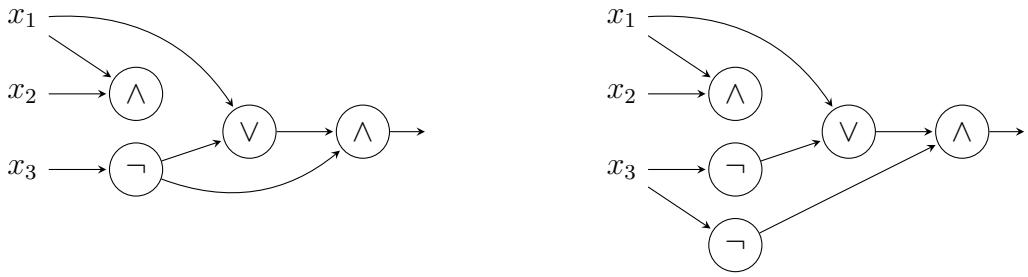


Рис. 12.2. Булева схема не задающая формула и её преобразование к формуле

Полнота базисов и классов функций. Замыкание классов

Вернёмся к вопросу, который мы ранее поставили в терминах формул: как понять, реализуют ли схемы с базисом B все булевы функции, т.е. представима ли любая булева функция f в схеме с базисом B. Базис, обладающий этим свойством называется ***полным***. Построив разложение в ДНФ произвольной булевой функции, мы доказали, что стандартный базис $\{\wedge, \vee, \neg\}$ полный. Используя законы Моргана легко показать, что полными базисами являются также $\{\wedge, \neg\}$ и $\{\vee, \neg\}$. Ниже, при изучении монотонных функций мы покажем, что базис $\{\wedge, \vee\}$ не является полным.

Полноту базиса B доказать очень легко: достаточно выразить функции некоторого полного базиса, например стандартного, через функции из B. Докозательство неполноты требует некоторой изобретательности, пример которой мы продемонстрируем ниже. Теорема Поста, формулировкой которой мы закончим лекцию позволяет легко проверить полноту любого (конечного) базиса.

Понятие полноты тривиальным образом переносится на классы булевых функций — подмножества всех булевых функций. Класс функций является ***полным***, если совпадает с классом всех булевых функций. Тривиальность определения компенсируется произвольностью способа описания класса функций. Класс функций можно описать с помощью схем: зафиксируем базис B и определим $\mathrm{cl}(B)$ как класс функций, состоящий из функций, реализуемых схемами в базисе B. Заметим, что базис B сам является классом функций, поэтому операция cl называется ***замыканием***. А класс $\mathcal{F}$ называется ***замкнутым***, если $\mathcal{F} = \mathrm{cl}(\mathcal{F})$. Так, определим класс функций D как класс функций, представимых через ДНФ. Мы знаем, что $\mathrm{cl}(D) = D$ и D совпадает с классом всех булевых функций.

Упражнение 9. Является ли полным класс симметрических функций? Является ли он замкнутым?

12.3 Монотонные функции

Монотонность (произвольной, не обязательно булевой) функции определяется относительно некоторого (частичного) порядка. В случае монотонных функций в математическом анализе используется стандартный порядок $\leqslant$ на $\mathbb{R}$ а для булевых функций нам пригодится покоординатный порядок (определённый на предыдущей лекции) на булевых векторах длины n.

Напомним, что $\{0,1\}^n = \{0,1\} \times \{0,1\} \times \ldots \times \{0,1\}$ — множество двоичных слов (или двоичных векторов) длины n. Также напомним определение покоординатного порядка для булевых векторов $a, b \in \{0,1\}^n$: $a \leqslant b$ если для всех $i \in \{1, \ldots, n\}$ справедливо $a_i \leqslant b_i$. Таким образом, булева функция $f : \{0,1\}^n \to \{0,1\}$ ***монотонна***, если $\forall a, b \in \{0,1\}^n : a \leqslant b \Rightarrow f(a) \leqslant f(b)$.

Связь с подмножествами

Использовав биекцию между двоичными словами длины n и подмножествами n-элементного множества получаем, что каждой монотонной булевой функции $f : \{0,1\}^n \to \{0,1\}$ соответствет функция $F : 2^{\{1,\ldots,n\}} \to \{0,1\}$, которая определяет для каждого подмножества, является ли оно хорошим (1) или плохим (0) с сохранением монотонности: $A \subseteq B \Rightarrow F(A) \leqslant F(B)$. Эта связь важна для практики: числа от 1 до n нумеруют признаки, например, олимпиады абитуриентов, а дальше вуз определяет какой набор олимпиад достаточен для поступления без экзаменов.

Неполнота класса монотонных функций

Монотонные булевы функции образует важный класс булевых функций, который мы обозначим через M. Он состоит из функций, реализуемых схемами с базисом $\{\wedge, \vee, 1, 0\}$; доказательство этого утверждения мы оставили в качестве домашнего задания.

Упражнение 10. Докажите, что $\mathrm{cl}(\{\wedge, \vee, 1, 0\}) = M$.

Также класс M можно описать как класс, состоящий из функций-констант $0, 1$ и функций, реализуемых схемами в базисе $\{\wedge, \vee\}$.

Докажем, что базис $\{\wedge, \vee\}$ не полный.

Утверждение 14. *Любая схема в базисе $\{\wedge, \vee\}$ на входе из одних нулей возвращает 0.*

Доказательство. Докажем утверждение индукцией по числу присваиваний в схеме. В случае нуля присваиваний схема возвращает значение переменной, поэтому база выполняется. Если для схем глубины меньше n утверждение верно, то либо $s_n = s_i \vee s_j$, либо $s_n = s_i \wedge s_j$, $i, j < n$. В обоих случаях, по предложению индукции $s_i = s_j = 0$ и как результат, $s_n = 0$. □

Таким образом мы доказали, что ни одна монотонная булева функция не вычисляет отрицание, а потому класс M не полный.

Подсчёт монотонных булевых функций

Подсчёт числа монотонных булевых функций от n-переменных является открытой математической проблемой — точная формула (а точнее вычислимая за разумное время) неизвестна. Мы приведём нижнюю оценку на это число. Для этого нам пригодится интерпретировать булеву функцию как раскраску ориентированного булева куба OB_n в два цвета: вершина a красится в $f(a)$. Напомним, что $V(OB_n) = \{0,1\}^n$, а ребро ведёт от вершины a к вершине b, если a и b отличаются в единственной позиции i и $a_i = 0$, $b_i = 1$. Переведём определение монотонной функции на язык графов: функция является монотонной тогда и только тогда, когда на любом пути из вершины $(0,0,\ldots,0)$ в вершину $(1,1,\ldots,1)$ нет перехода цветов от 1 к 0. Отметим, что граф OB_n является диаграммой Хассе для покоординатного порядка на булевых векторах длины n, отсюда и берётся перевод определения на язык графов: если f не монотонна, то для некоторых векторов a и b, таких что $a \leqslant b$, выполняется $f(a) = 1$ и $f(b) = 0$; тогда в графе есть путь из a в b, а значит на этом пути есть ребро из вершины цвета 1 в вершину цвета 0.

Для получения оценки на число монотонных булевых функций мы опишем некоторые монотонные булевы функции и оценим их число. Нам потребуются вспомогательные определения. Обозначим через $|a|$ число единиц в булевом векторе a. Обозначим через $V_i = \{a : |a| = i, a \in \{0,1\}^n\}$, назовём V_i — i-м ***слоем*** булева куба. Будем считать, что $n = 2k$, тогда V_k — центральный слой булева куба. Легко видеть, что между вершинами одного слоя нет рёбер.

Перейдём к описанию подмножества монотонных булевых функций. Функция f определяется раскраской центрального слоя: $f(a) \in \{0,1\}$ при $|a| = k$, $f(a) = 0$ при $|a| < k$, $f(a) = 1$ при $|a| > k$. Легко видеть, что число таких функций равно $2^{|V_k|}$, таким образом число всех монотонных функций не меньше $2^{|V_k|}$. Осталось оценить мощность множества V_k: число двоичных слов длины $2k$, в которых ровно половина единиц, равно центральному биномиальному коэффициенту $\binom{2k}{k}$, для которого мы ранее получили оценкцу $\binom{2k}{k} \geqslant \frac{2^{2k}}{2k+1}$. Таким образом, число монотонных булевых функций от $2k$ переменных не меньше чем

$$2^{\frac{2^{2k}}{2k+1}}.$$

12.4 Многочлены Жегалкина

Формулы в базисе $\{1, \oplus, \wedge\}$ выглядят как многочлены:

$$1 \oplus x_1x_2 \oplus x_1x_3 \oplus x_1x_2x_3 \tag{4}$$

И такие многочлены и сам базис носят имя Ивана Ивановича Жегалкина. Если вы будете изучать дальше высшую алгебру, то узнаете, что с алгебраической точки зрения это многочлены над полем F_2. В случае если функция задана формулой в базисе Жегалкина, можно, раскрыв все скобки и приведя подобные слагаемые, привести многочлен к ***стандартному виду***. Опишем многочлен Жегалкина в стандартном виде явно. Отметим, что в случае «сложения» двух одинаковых одночленов, они взаимноуничтожаются; каждый одночлен состоит из конъюнкции

переменных, поэтому определён множеством $S \subseteq \{1, 2, \ldots, n\}$:

$$S \mapsto \bigwedge_{i \in S} x_i \tag{5}$$

Так множеству $\{1, 2, 3\}$ соответствует конъюнкт $x_1 x_2 x_3$. Каждый конъюнкт (5) либо входит в многочлен Жегалкина, либо нет. Обозначим через $\mathcal{F} \subseteq 2^{\{1,\ldots,n\}}$ множество из всех S, конъюнкты для которых входят в многочлен; конъюнктом для $S = \varnothing$ является 1. Так для многочлена (4) $\mathcal{F} = \{\varnothing, \{1, 2\}, \{1, 3\}, \{1, 2, 3\}\}$.

Таким образом, многочлен Жегалкина в стандартном виде однозначно определяется через $\mathcal{F}$ как

$$\bigoplus_{S \in \mathcal{F}} \bigwedge_{i \in S} x_i \tag{6}$$

Обозначим через g функцию, которая по $\mathcal{F}$ возвращает булеву функцию $f(x_1, \ldots, x_n)$, заданную многочленом Жегалкина в стандартном виде (6). Легко проверить, что базис Жегалкина является полным: $1 \oplus x = \neg x$, а базис $\{\neg, \wedge\}$ является полным. Значит, для любой булевой функции f существует реализующий её многочлен Жегалкина, а поскольку любой многочлен Жегалкина приводится к стандартному виду, получаем, что g — сюръекция (для каждой булевой функции $f(x_1, \ldots, x_n)$ существует класс $\mathcal{F}$, такой что $g(\mathcal{F}) = f$).

Подсчитаем теперь число различных $\mathcal{F}$ и булевых функций n переменных: оба числа равны 2^{2^n}, отсюда и из сюръективности g получаем, что g — биекция, то есть разным булевым функциям соответствуют разные многочлены Жегалкина в стандартном виде.

Теорема Поста*

Теорема Поста является одним из важных результатов в области булевых функций. Её доказательство достаточно техническое, мы разбили его на задачи в классных и домашних листках; также с ним можно ознакомиться в [2; 9]. Для её формулировки нам потребуется определить ***классы Поста***:

M монотонные функции

L ***линейные функции***: $L = \mathrm{cl}(\{\oplus, 1\})$

T_0 ***функции, сохраняющие ноль*** $f \in T_0 \iff f(0, 0, \ldots, 0) = 0$

T_1 ***функции, сохраняющие единицу*** $f \in T_1 \iff f(1, 1, \ldots, 1) = 1$

S ***самодвойственные функции*** $f \in S \iff \overline{f(\overline{x_1}, \overline{x_2}, \ldots, \overline{x_n})} = f(x_1, x_2, \ldots, x_n)$

Все эти классы являются неполными.

Теорема 10 (теорема Поста)**.** *Замкнутый класс булевых функций $\mathcal{F}$ является неполным тогда и только тогда, когда содержится в одном из классов Поста.*

Из теоремы Поста следует, что если добавить к любому классу Поста хотя бы одну функцию, которая в нём не лежала, и взять замыкание, то получится полный класс. Классы обладающие таким свойством называются ***предполными***. Также Пост доказал, что только классы M, L, T_0, T_1, S являются предполными.

Лекция 13

Производящие функции I

План:

1. Определения и примеры
   - Производящая функция бинома Ньютона
     - Сила дифференцирования: $\sum\limits_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$
     - $(1+x)^m(1+x)^n = (1+x)^{m+n}$. Отсюда следует

$$\binom{m+n}{k} = \binom{m}{0}\binom{n}{k} + \binom{m}{1}\binom{n}{k-1} + \ldots + \binom{m}{k}\binom{n}{0}.$$

   - Аналитическое представление производящей функции:

$$\frac{1}{1-x} = 1 + x + x^2 + \ldots$$

2. Применение для решения комбинаторных задач
   - Задача Муавра
   - Задача о счастливых билетах
   - Найти число целочисленных решений системы

$$\begin{cases} x_1 + x_2 + x_3 = 40, \\ 4 \leqslant x_1 \leqslant 15, \\ 9 \leqslant x_2 \leqslant 18, \\ 5 \leqslant x_3 \leqslant 16. \end{cases}$$

Производящие функции — важный инструмент в комбинаторике. Они открывают новый взгляд на задачи о подсчёте комбинаторных объектов. Прежде чем давать определения, начнём с примера. Вспомним, что по биному Ньютона

$$(1+x)^n = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \ldots + \binom{n}{k}x^k + \ldots \binom{n}{n}x^n.$$

Продифференцировав обе части, получим

$$n(1+x)^{n-1} = \binom{n}{1} + 2\binom{n}{2}x + \ldots + k\binom{n}{k}x^{k-1} + \ldots n\binom{n}{n}x^{n-1}.$$

Подставив $x = 1$ получим формулу

$$n \times 2^{n-1} = \sum_{k=1}^{n} k\binom{n}{k},$$

которую мы уже доказывали комбинаторно (подсчитав сколькими способами можно выбрать отряд солдат с командиром или послать делегацию студентов с лидером на конференцию). Но здесь мы не пользовались комбинаторным доказательством, а использовали свойство дифференцирования!

Этот пример показывает, что удобно работать с последовательностью комбинаторных чисел, представив их многочленами. Если последовательность бесконечная (как, например, числа Фибоначчи), то и многочлен получится «бесконечным», такие «бесконечные» многочлены называют формальными степенными рядами. Их изучению отводится место во втором семестре курса математического анализа (ряды Тейлора), однако для наших нужд хватит знакомства с формулой Тейлора (а также навыков дифференцирования и интегрирования).

Дадим определение производящей функции. ***Производящей функцией*** конечной последовательности $a_0, a_1, a_2, \ldots, a_n$ называется многочлен

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n;$$

производящей функции бесконечной последовательности $a_0, a_1, a_2, \ldots, a_n, \ldots$ называется формальный степенной ряд

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots.$$

Мы будем для удобства считать, что конечная последовательность — частный случай бесконечной, у которой начиная с некоторого номера все члены нулевые. Так мы воспользуемся тем, что многочлен — частный случай ряда, и можем проводить рассуждения только для рядов.

Читатель может удивиться тому, что производящая функция — не функция. Этому казусу есть следующее объяснение. В случае конечной последовательности, многочлен является функцией и тут проблем нет. А бесконечные последовательности удобо представлять в виде ряда Тейлора аналитической функции. Пусть функция $f(x)$ имеет производную любого порядка в точке 0. Рядом Тейлора (в точке 0) функции f называется формальный степенной ряд

$$f(0) + f'(0)x + \frac{f''(x)}{2!}x^2 + \ldots = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!}x^n.$$

Пример 26. Из математического анализа известно, что n-ая производная функции $\frac{1}{1-x}$ в нуле равна 1. Таким образом

$$\frac{1}{1-x} = 1 + x + x^2 + \dots$$

В случае если ряд Тейлора аналитической функции[1] f совпадает с производящей функцией последовательности $\{a_n\}_{n=0}^{\infty}$, то мы также называем f производящей функцией последовательности $\{a_n\}_{n=0}^{\infty}$ и говорим, что производящая функция ***задана аналитически***.

Утверждение 15. *Пусть $f(x)$ и $g(x)$ — аналитические производящие функции последовательностей $\{a_n\}_{n=0}^{\infty}$ и $\{b_n\}_{n=0}^{\infty}$ соответственно. Из определения ряда Тейлора и свойств дифференцирования следует, что*

- $f(x) + g(x)$ *— производящая функция последовательностей* $\{a_n + b_n\}_{n=0}^{\infty}$;
- $f'(x)$ *— производящая функция последовательностей* $\{(n+1)a_{n+1}\}_{n=0}^{\infty}$;
- $xf(x)$ *— производящая функция последовательности* $0, a_0, a_1, \dots$;
- $f(0) = a_0$.

Идея доказательства. Разложим функции f и g по формуле Тейлора в точке 0 до n-го порядка: первые n членов формулы Тейлора совпадают с первыми n членами ряда Тейлора. Убедившись, что свойства верны для формул Тейлора (соответствующих начальному отрезку последовательности), получим, что они верны и для рядов. □

Перейдём к применению техники производящих функций.

Пример 27. Докажем тождество

$$\binom{m+n}{k} = \sum_{j=0}^{k} \binom{m}{j}\binom{n}{k-j}.$$

Для этого выразим через биномы Ньютона $(1+x)^m$ и $(1+x)^n$, и воспользуемся фактом

$$(1+x)^{m+n} = (1+x)^m(1+x)^n;$$

$$\sum_{k=0}^{m+n} \binom{m+n}{k} x^k = \left(\sum_{j=0}^{m} \binom{m}{j} x^j\right)\left(\sum_{i=0}^{n} \binom{n}{i} x^i\right).$$

Из разложений по биному получаем, что $\binom{m+n}{k}$ — коэффициент при одночлене x^k в левой части равенства, а $\sum_{j=0}^{k} \binom{m}{j}\binom{n}{k-j}$ — коэффициент при x^k в правой части.

[1] Для наших нужд хватит неформального понимания «функция задана формулой». Функция называется аналитической в точке 0, если её ряд Тейлора сходится к f в некоторой окрестности нуля, т.е. значение вычисленное через ряд в точке x в этой окрестности совпадает с $f(x)$. Все достаточно простые функции (а только такие у нас и возникнут) аналитические в нуле. Для наших нужд хватит даже существования у функции производной любого порядка в нуле.

Действительно, зафиксируем x^j из первого множителя в правой части равенства, коэффициент при этом члене $\binom{m}{j}$; чтобы получить x^k нужно умножить x^j на x^{k-j} из второй скобки, а коэффициент при x^{k-j} равен $\binom{n}{k-j}$. □

Пример 28. Задача Муавра: найти число решений уравнения $x_1 + x_2 + x_3 = 11$ в неотрицательных целых числах.

Мы уже умеем решать эту задачу комбинаторно, но сейчас мы разберём решение через производящие функции, чтобы потом обобщить эту технику на более сложный случай. Глядя на условие, сходу неясно, как применить этот метод, ведь в условии не фигурирует последовательность. Решим более общую задачу:

$$x_1 + x_2 + x_3 = n;$$

Последовательность $\{a_n\}$ состоит из числа решений уравнения с фиксированным n. Вспомним, что

$$\frac{1}{(1-x)} = 1 + x + x^2 + \dots$$

и рассмотрим выражение

$$\frac{1}{(1-x_1)}\frac{1}{(1-x_2)}\frac{1}{(1-x_3)} = (1 + x_1 + x_1^2 + \dots)(1 + x_2 + x_2^2 + \dots)(1 + x_3 + x_3^2 + \dots).$$

Раскроем мысленно скобки в правой части и обнаружим, что каждый одночлен степени n соответствует решению задачи, а именно одночлену $x_1^{k_1}x_2^{k_2}x_3^{k_3}$, в котором $k_1 + k_2 + k_3 = n$ соответствует решение $x_i = k_i$. Ясно, что это соответствие взаимно однозначно — для каждого набора (k_1, k_2, k_3) найдётся единственный соответствующий одночлен. Таким образом число решений задачи Муавра есть число одночленов степени n в произведении рядов Тейлора.

С ответом в такой форме работать неудобно, поэтому отметим, что если мы сотрём индексы у переменных (или, что то же самое, подставим $x_1 = x_2 = x_3 = x$), то переменная останется одна, а ответом теперь будет просто коэффициент при x^n. Таким образом, производящей функцией искомой последовательности является функция

$$\frac{1}{(1-x)^3},$$

а ответом на вопрос задачи будет коэффициент при x^{11}.

В общем случае, производящей функцией для задачи Муавра

$$x_1 + x_2 + \dots + x_m = n$$

будет

$$\frac{1}{(1-x)^m}.$$

□

Для получения численного ответа из примера нам потребуются факты из математического анализа. Для произвольного числа α обозначим

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1) \times \dots \times (\alpha-k+1)}{k!}.$$

Известно, что

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n.$$

Эту формулу как раз и доказал Ньютон, а не бином, известный и до него.

Для $(1-x)^m$ получим

$$(1-x)^m = \sum_{n=0}^{\infty} (-1)^n \binom{-m}{n} x^n$$

и преобразуем

$$\begin{aligned}(-1)^n \binom{-m}{n} &= (-1)^n \frac{-m(-m-1)\times\ldots\times(-m-n+2)(-m-n+1)}{n!} \\ &= \frac{(m+n-1)(m+n-2)\times\ldots\times(m+1)m}{n!} = \binom{m-1+n}{n}.\end{aligned}$$

В предпоследнем переходе мы домножили каждую скобку на (-1) — число скобок равно n, как и число.

Таким образом

$$(1-x)^{-m} = \sum_{n=0}^{\infty} \binom{m-1+n}{n} x^n.$$

Легко убедиться, что полученное решение совпадает с тем, которое мы получили методом точек и перегородок.

Пример 29. Найти число целочисленных решений системы

$$\begin{cases} x_1 + x_2 + x_3 = 40, \\ 4 \leqslant x_1 \leqslant 15, \\ 9 \leqslant x_2 \leqslant 18, \\ 5 \leqslant x_3 \leqslant 16. \end{cases}$$

Мы уже решали задачи Муавра с ограничениями, и использовали для этого метод включений-исключений, что приводило к громоздким выкладкам. Метод производящих функций позволяет решать задачи Муавра с ограничениями проще и изящнее. Как и в предыдущем примере, найдём производящую функцию последовательности $x_1 + x_2 + x_3 = n$ (с учётом ограничений), из которой нам потребуется коэффициент при x^{40}. Следуя тем же шагам, сначала запишем многочлен многих переменных, в котором каждый одночлен кодирует решение:

$$\begin{aligned}&(x_1^4 + x_1^5 + \ldots + x_1^{15}) \times (x_2^9 + x_2^{10} + \ldots + x_2^{18}) \times (x_3^5 + x_3^6 + \ldots + x_3^{16}) = \\ &= x_1^4(1 + x_1 + \ldots + x_1^{11}) \times x_2^9(1 + x_2 + \ldots + x_2^9) \times x_3^5(1 + x_3 + \ldots + x_3^{11}).\end{aligned}$$

Для последующих преобразований воспользуемся формулой суммы геометрической прогрессии:

$$1 + x + \ldots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

Получим

$$x_1^4 x_2^9 x_3^5 \times \frac{1-x_1^{12}}{1-x_1} \times \frac{1-x_2^{10}}{1-x_2} \times \frac{1-x_3^{12}}{1-x_3}.$$

Отождествив переменные получим искомую производящую функции:

$$x^{18} \times \frac{1-x^{12}}{1-x} \times \frac{1-x^{10}}{1-x} \times \frac{1-x^{12}}{1-x}.$$

Ясно, что коэффициент при x^{40} этой производящей функции равен коэффициенту при x^{22} следующей производящей функции:

$$\frac{(1-x^{12})^2 (1-x^{10})}{(1-x)^3} = \left(1-2x^{12}+x^{24}\right)(1-x^{10})(1-x)^{-3}.$$

Раскроим скобки и избавимся от членов выше 22-ой степени, поскольку они не вносят вклад в искомый коэффициент:

$$\left(1-x^{10}-2x^{12}+2x^{22}\right)\sum_{k=0}^{22}(-1)^k\binom{-3}{k}x^k =$$
$$= \binom{-3}{22}x^{22} - \binom{-3}{12}x^{22} - 2\binom{-3}{12}x^{22} + x^{22} + (\ldots),$$

где в $(\ldots)$ оставлены все одночлены не двадцать второй степени.

Таким образом, получаем ответ

$$\binom{-3}{22} - \binom{-3}{12} - 2\binom{-3}{10} + 2,$$

Который можно преобразовать

$$\binom{24}{22} - \binom{14}{12} - 2\binom{12}{10} + 1 = \binom{24}{2} - \binom{14}{2} - 2\binom{12}{2} + 2 = 55.$$

□

Пример 30. Задача о счастливых билетах. Найти число последовательностей $(a_1, a_2, a_3, b_1, b_2, b_3)$, $a_i, b_i \in \{0, \ldots, 9\}$, таких что $a_1 + a_2 + a_3 = b_1 + b_2 + b_3$.

Найдём сначала производящую функцию $C(x)$ последовательности

$$a_1 + a_2 + a_3 = n \ (a_i \in \{0, \ldots, 9\}).$$

Это задача Муавра с ограничениями; получаем, что

$$C(x) = \left(1 + x + x^2 + \ldots + x^9\right)^3 = 1 + c_1 x + c_2 x^2 + \ldots + c_{27} x^{27}.$$

Коэффициент c_n равен числу способов разложить число n в сумму $a_1 + a_2 + a_3$ и это число совпадает с числом разложений $n = b_1 + b_2 + b_3$. Таким образом, число счастливых билетов, в которых $a_1 + a_2 + a_3 = n$, равно c_n^2 по правилу произведения. Отсюда, получаем, что общее число счастливых билетов равно

$$\sum_{n=0}^{27} c_n^2.$$

Список литературы

1. Лекции по дискретной математике / М. Вялый, В. Подольский, А. Рубцов, Д. Шварц, А. Шень. — ИД НИУ ВШЭ. Черновик: https://publications.hse.ru/mirror/pubs/share/direct/393719078.pdf, 2021.
2. *Журавлёв Ю. И.*, *Флёров Ю. А.*, *Федько О. С.* Дискретный Анализ. Комбинаторика. Алгебра логики. Теория графов. — М.: МФТИ, 2012.
3. *Биркгоф Г.*, *Барти Т.* Современная прикладная алгебра. — Издательство "Мир", 1976.
4. *Lehman E.*, *Leighton F. T.*, *Meyer A. R.* Mathematics for Computer Science. — United Kingdom : Samurai Media Limited, 2017. — URL: https://courses.csail.mit.edu/6.042/spring17/mcs.pdf.
5. *Sipser M.* Introduction to the Theory of Computation. — Third. — Boston, MA : Course Technology, 2013. — ISBN 113318779X.
6. *Мендельсон Э.* Введение в математическую логику: — УРСС, 2010. — ISBN 9785397013871.
7. *Дистель Р.* Теория графов. — Новосибирск: Институт математики, 2002.
8. *Lovasz L.*, *Vesztergombi K.* Discrete Mathematics. Lecture Notes, Yale University. — 1999. — URL: http://www.cs.elte.hu/~lovasz/dmbook.ps.
9. *Яблонский С.* Введение в дискретную математику. — М.: Высшая школа, 2003.
10. *Зуев Ю.* По океану дискретной математике: От перечислительной комбинаторики до современной криптографии. Т.1: Основные структуры. Методы перечисления. Булевы функции. — М.: Книжный дом «ЛИБРОКОМ», 2012.
11. *Зуев Ю.* По океану дискретной математике: От перечислительной комбинаторики до современной криптографии. Т.2: Графы. Алгоритмы. Коды, блоксхемы, шифры. — М.: Книжный дом «ЛИБРОКОМ», 2012.
12. *Грэхем Р.*, *Кнут Д.*, *Паташник О.* Конкретная математика. Математические основы информатики. — Вильямс, 2010.
13. *Харари Ф.* Теория графов. Изд. 2-е. — М.: Эдиториал УРРС, 2003.
14. *Андерсон Д. А.* Дискретная математика и комбинаторика. — М.: Вильямс, 2003.

15. Сборник задач по дискретному анализу. Комбинаторика. Элементы алгебры логики. Теория графов / Ю. И. Журавлёв, Ю. А. Флёров, О. С. Федько, Т. М. Дадашев. — М.: МФТИ, 2004.